BEST AVAILABLE COPY

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 016 314 A1

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 80100374.0

(22) Date of filing: 24.01.80

(51) Int. Cl.³: A 63 F 9/22
G 09 B 7/04, H 04 N 5/00

(30) Priority: 05.02.79 US 9533

(43) Date of publication of application: 01.10.80 Bulletin 80/20

(84) Designated Contracting States: DE FR GB IT NL SE

(71) Applicant: Best, Robert MacAndrew
16016 - 9th Avenue N.E.
Seattle, Washington 98155(US)

(72) Inventor: Best, Robert MacAndrew
16016 - 9th Avenue N.E.
Seattle, Washington 98155(US)

(74) Representative: Patentanwälte Dipl.-Ing. A. Grünecker, Dr.-Ing. H. Kinkeldey, Dr.-Ing. W. Stockmair, Dr. rer. nat. K. Schumann, Dipl.-Ing. P.H. Jakob, Dr. rer. nat. G. Bezold Maximilianstrasse 43
D-8000 München 22(DE)

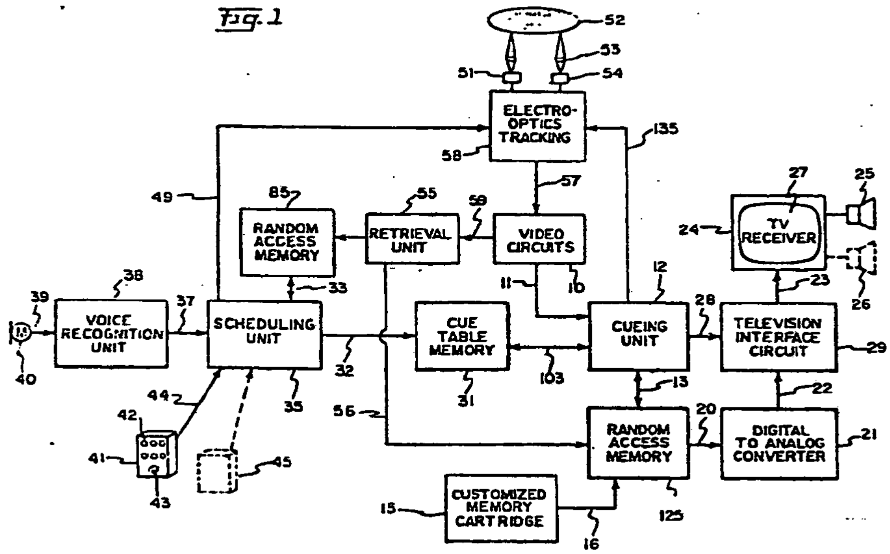
(54) **Method and apparatus for voice dialogue between a video picture and a human.**

(57) A video entertainment system by which human viewers conduct simulated voice conversations with screen actors in a prerecorded branching movie shown on a television screen (27). A voice-recognition unit (38) recognizes a few words spoken by a viewer at branch points in the movie. A different set of words may be used at each branch point. A hand-held unit (41) displays prompting messages to inform each viewer of the words that can be recognized at each branch point. A scheduling unit (35) assembles cueing commands specifying which video frames, cartoon frames, messages, and audio portions are to be presented at which instant of time. A cueing unit (12) executes these commands by generating precisely timed video and audio signals, so that a motion picture with lip-synchronized sound is presented to the viewer who vocally influences the course of the movie.

EP 0 016 314 A1

./...

Fig. 1


52
53
51
54
ELECTRO-OPTICS TRACKING
58
135
57
49
85
RANDOM ACCESS MEMORY
55
RETRIEVAL UNIT
59
VIDEO CIRCUITS
27
25
24
TV RECEIVER
10
11
12
23
26
38
39
VOICE RECOGNITION UNIT
37
SCHEDULING UNIT
33
CUE TABLE MEMORY
CUEING UNIT
28
TELEVISION INTERFACE CIRCUIT
29
40
32
103
31
35
13
22
44
42
56
20
RANDOM ACCESS MEMORY
DIGITAL TO ANALOG CONVERTER
21
41
45
43
15
CUSTOMIZED MEMORY CARTRIDGE
16
125

TITLE MODIFIED
see front page

TECHNICAL FIELD OF THE INVENTION

This invention relates to video systems, voice-recognition devices, branching movies, and picture/sound synchronization.

BACKGROUND OF THE INVENTION

While watching a prior-art sound movie, a viewer often experiences a vicarious sense of involvement. But the viewer cannot actively participate in the movie, because the viewer cannot talk with the screen actors and have them reply responsively. Applying prior-art voice-recognition techniques to control prior-art branching movies would not provide a natural conversational dialog because of the following problem: If the number of words which a viewer of any age and sex can speak and be understood by the apparatus is sufficiently large to permit a natural conversation, then prior-art voice-recognition techniques are unreliable. Conversely, if the number of words is restricted to only a few words to make voice recognition reliable, then natural conversation would not result. It is also necessary for the picture to be responsive to a viewer's voice and be synchronized with the spoken reply. These problems are not addressed in the prior art.

An example of a prior-art branching movie is shown in U.S. Patent 3,960,380, titled "Light Ray Gun and Target Changing Projectors". This system uses a pair of film projectors which present two alternatives (hit or miss) at each branch point. An example of a prior-art device for interactive voice dialog is shown in U.S. Patent 4,016,540. This device does not present a motion picture. An example of a prior-art video editing system is shown in U.S. Patent 3,721,757. This system displays a sequence of video excerpts specified by a control program of stored videodisc addresses which comprise a sequential (not branching) movie. To change this sequence the editor alters the program. In the present invention the viewer does not alter the program.

SUMMARY OF THE INVENTION

This invention is a video entertainment system by which one or more human viewers influence the course of a prerecorded movie and conduct a simulated two-way voice conversation with screen actors in the movie. Viewers talk to the screen actors and the actors talk back and talk responsively. The system thus provides an illusion of individualized and active participation in the movie as if each viewer were a participant in a real-life drama.

To simplify performing of voice-recognition on the voices of different viewers, regardless of age and sex, while at the same time using a large vocabulary of computer-recognizable words, the words to be recognized at each branch point in the movie are restricted to two or a few words such as "yes" and "attack". The words which a viewer may use at each branch point will often be different from words used at other branch points. The apparatus informs the viewers of what words they can use by displaying prompting messages on a hand-held display device and/or with prompting words spoken by a screen actor. The display device also contains a microphone into which a viewer speaks a selected word at each branch point.

To permit a viewer to ask questions or make other remarks which require whole sentences, at some branch points the hand-held device displays a list of appropriate sentences. Next to each sentence is a push button which when pressed causes a voice recording of the displayed sentence to be played or synthesized. A screen actor then responds to the sentence or question as if the viewer had spoken it. The voice recording is selected from several recordings of different voices so that the played voice most closely resembles the voice of the viewer. Recordings of the viewers' names are inserted into the audio so that the actors speak to each viewer

using the viewer's own name.

A precisely timed sequence of video frames and lip-synchronized audio is generated for each story line according to a prerecorded schedule of control commands which is continually updated as the movie proceeds and alternative branches in the movie are chosen. When each viewer/player makes a choice at a branch point in the movie, a new set of commands is scheduled for execution.

These commands are of two kinds: story commands which define a branching structure of possible alternative story lines, and cueing commands which specify timing of video frames and audio portions. At each branch point in a network of story commands, two or more story commands may point to alternative chains or branching structures of story commands representing alternative sequences of scenes in the movie.

A scheduling unit processes a chain of story commands and assembles a schedule of cueing commands specifying precisely which video frames, cartoon frames, and portions of audio are to be presented at which instant of time. A cueing unit executes these commands by generating precisely timed video and audio signals, so that a movie with lip-synchronized sound is presented to the viewer.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram of a special-purpose microcomputer which processes prerecorded video frames.

Fig. 2 is a block diagram of a special-purpose microcomputer which processes digitally generated animated cartoons.

Fig. 3 is a detailed block diagram of scheduling unit 35.

Fig. 4 is a detailed block diagram of cueing unit 12.

Fig. 5 illustrates a data structure network of story commands.

Fig. 6 is a detailed block diagram of control circuits 62.

Fig. 7 illustrates how two different audio signals may be synchronized with a common set of multiple-use video frames.

Fig. 8 is a block diagram of one type of mixer 129 for digitized audio.

Fig. 9 is a program flowchart for scheduling unit 35.

Fig. 10 is a cartoon illustrating a branch point in a movie when a viewer may cause alternative story lines by speaking into a microphone.

Fig. 11 is a storyboard diagram illustrating one episode of a branching movie.

Fig. 12 is a detailed block diagram of initiator switching unit 131 combined with terminator switching unit 118.

Fig. 13 is a pictorial view of a hand-held input device by which a viewer may influence the course of a branching movie.

Fig. 14 is a program flowchart for cueing unit 12.

Fig. 15 is a continuation of Fig. 14 for video cue commands.

Fig. 16 is a continuation of Fig. 14 for audio cue commands.

DETAILED DESCRIPTION OF THE INVENTION

Referring to Fig. 1, in one embodiment of this invention, a special-purpose microcomputer, which includes units 35, 55, 12 and other units, is connected to a conventional television receiver 24 and to a conventional random-access videodisc reader which includes unit 58 for automatic seeking of track addresses and for automatic tracking of disc tracks. One or more hand-held input units 41, each containing a microphone 40, are also connected to the microcomputer by wire 44 or by wireless transmission using transceiver 171 (Fig. 2).

The microcomputer in Figs. 1 and 2 controls reading of information from videodisc 52 and processes the viewer's inputs from microphone 40. Cartridge 15 containing digitized recordings of the viewer's names may plug into the microcomputer.

The microcomputer shown in Fig. 1 includes: voice recognition unit 38, scheduling unit 35, retrieval unit 55, and cueing unit 12. The microcomputer also contains conventional random access memories 31, 85 and 125, digital-to-analog converter 21 to generate audio signal 22, conventional RF-modulator circuit 29 to interface with the television receiver 24, and prior-art video circuits 10 for vertical/horizontal sync separation, demodulation, chroma separation and phase invertion.

Unit 58 may be one or more conventional videodisc tracking units, such as the one described in U.S. Patent 4,106,058. Two optical read heads 51 and 54 and 2-channel circuitry in unit 58 are preferred for reading camera-originated video frames so that one read head 54 can be moving to the next position on the disc while the other read head 51 is reading the current video frames or vice versa. One channel is sufficient for embodiments which use digitally generated animated cartoons in lieu of camera-originated video.

The video signal for each frame passes from tracking unit 58 through video circuit 10, cueing unit 12 and interface circuit 29 to television receiver 24. Digitized audio passes from video circuit 10 through retrieval unit 55, memory 125, digital to analog converter 21, and interface circuit 29 to television receiver 24. The control commands pass from circuit 10 through retrieval unit 55, memory 85, scheduling unit 35, memory 31, to cueing unit 12. Memories 85, 86 and 125 may be different portions of a common memory, but are shown separately in the drawings for clarity.

Retrieval unit 55 is a conventional peripheral input

controller which stores into memory the digitally coded blocks of information obtained from videodisc 52. This information includes control data (cue commands and story commands) which unit 55 stores into memory 85 (and memory 86 in Fig. 3) for use by scheduling unit 35, and compressed audio and/or graphics data which unit 55 stores into memory 125 via line 56 for use by cueing unit 12.

Scheduling unit 35 (the circuit detailed in Fig. 3 or a microprocessor programmed to perform equivalent functions) is the master scheduler and has final control of the course of the movie. By way of example, Fig. 9 illustrates a program for performing the main functions of scheduling unit 35. Scheduling unit 35 may request successive blocks of control information from retrieval unit 55 and output into random access memory 31 a schedule (called a cue table) of tasks for cueing unit 12 to do. Scheduler 35 repeatedly updates the cue table schedule as the movie progresses. Scheduler 35 processes the choices of the human viewers which are input through one or more hand-held input units 41 and/or 45, and stores different commands into cue table 31 depending on the viewer's choices.

Cueing unit 12 (the circuit detailed in Fig. 4 or a microprocessor programmed to perform equivalent functions) repeatedly scans cue table 31 to get commands telling it what to do and the instant of time it should do it. By way of example, Figs. 14-16 illustrate a program for performing cueing unit 12 functions. Cueing unit 12 edits digitized audio and other data already stored in random access memory 125 by retrieval unit 55. This editing process is directed by the commands in cue table 31 and generates a continuous sequence of output records (into register 19 in Fig. 4) containing edited, mixed, and synchronized audio in compressed digital form. Some of these edited records may contain graphics

information (representing text, animation data, and/or special patterns) which are passed in cueing unit 12 to the graphics generator (block 126 in Fig. 4) which generates the video signals on line 146 representing the graphics display. This may consist of alphabetic characters which form titles large enough to be read from television screen 27, lines which form patterns, special shapes commonly found in video games, and/or animated cartoons.

Cueing unit 12 also controls the position of read head 51 or 54 which is currently reading video, and processes the composite video signal on line 11 from circuit 10. Although there may be many sequences of frames which occupy consecutive tracks on disc 52 (either spiral or circular), in general there will be frequent jumps to non-adjacent tracks. This random-access movement is controlled in a conventional manner by electro-optical tracking unit 58 using track address searching during vertical blanking intervals. If a large jump to a distant track address is required, the other read head is positioned by cueing unit 12 in response to a command in cue table 31 to move to the distant track, well in advance of the time it is needed, so that a switch to the other head may be made automatically (by block 142 in Fig. 4) during the scheduled vertical interval without a discontinuity in the picture.

The sequence in which tracks are accessed by each read head is specified by the commands in cue table 31. During picture intervals, cueing unit 12 scans cue table 31 for the next command or commands which specify the next track address required by each head.

In an alternative embodiment of this invention shown in Fig. 2, the data read from disc 52 and/or magnetic bubble memory 173 may be compressed binary data from which graphics generator 126 generates animated cartoons. The compressed data required to generate a cartoon.

frame may be stored in a fraction of a disc track. Alternatively, disc 52 and tracking unit 58 may be eliminated if magnetic bubble memory 173 has enough capacity to store the compressed data.

HOW THE INVENTION IS USED

At frequent branch points in the movie the apparatus presents the viewer with two or more alternatives to choose among, predetermined remarks to make to the actors, predetermined questions to ask, or the opportunity to change the course of the action or dialog.

Fig. 10 illustrates a typical branch point which leads either to a fight scene or a chase scene depending on the viewer's choice. In this illustration a chase will result. The video frames for the fight scene need not be wasted. They may be used in a later episode. Multiple choices are presented to the viewer in a sequence determined by previous choices. These may be displayed as titles on screen 27 or unit 41, or may be inferred by the viewers from the situation, or may be spoken by an actor. Such an actor, shown on the screen in Fig. 10 keeps the viewer or viewers informed on what is happening, what problems require a decision, what the alternatives are, and executes some of the actions selected by a viewer. Such an actor or actors guide the viewers into scenes which the videodisc recording is capable of providing.

The alternative words which are acceptable to the apparatus at each branch point may be explicitly spelled out for the viewer on a readable display such as the liquid crystal display 174 illustrated in Fig. 13. Each set of alternative words or phrases which the viewer may speak and be understood by voice recognition unit 38 may be sent to each unit 41 by transceiver 171. Cueing each display is performed by cueing unit 12 directed by a cue command. One cue command is scheduled for each message by scheduler 35. The messages are retrieved from

videodisc 52, or bubble memory 173 by retrieval unit 55 together with digitized audio and control commands.

The words displayed on display 174 may be the same as words spoken by a screen actor, or may be a subset of the words spoken by a screen actor, or may be additional words. These displayed words are the alternative responses which a viewer selects among and may include words used in lieu of whole phrases suggested by a screen actor. Words not suggested by the screen actor may also be displayed on display 174 to indicate alternative responses for speaking by a viewer. Indicator lamp 43 may be used to tell a viewer when a spoken response is expected and when a push-button response is expected.

So that a viewer may ask questions or make other remarks which are not responses to suggestions by a screen actor, multiple sentences may be displayed as a list on display 174 adjacent to corresponding push buttons 42 or touch pads. When a button 42 is pushed, a sound recording of a voice speaking the selected sentence is played through speaker 25 as a substitute for the viewer's part of the conversation. The screen actor then "responds" as if the words in the sound recording had been spoken by the viewer. Because the viewer selects the words which are actually sounded, the viewer will quickly learn to disregard the fact that the words have been put in his mouth. Pushing a button 42 selects both a simulated verbal response to the previous scene and also a new scene which corresponds to the simulated verbal response displayed on display 174. The selected scene includes the face and voice of the actor speaking words which are responsive to the viewer's selected verbal response.

To preserve naturalness and differences in age and sex, several alternative voices, all speaking the same words, may be recorded on disc 52 or in cartridge 15 together with corresponding story commands which are

UU16514

processed by scheduler 35 at such branch points. Scheduler 35 then schedules cue commands which point to the digitized sound recording having preanalyzed voice characteristics which most closely resemble the characteristics of the viewer's voice as determined by voice recognition unit 38.

Planning a branching movie is more complex than a conventional movie. Fig. 11 shows a simplified storyboard in which the rectangles represent conventional scenes and the ovals represent branch points. Note that multiple story branches (represented by arrows) can converge on a common scene. Chase scene 464 for example, can follow either branch point 462 or branch point 463 depending on an earlier choice at branch point 461. Branch points such as 466 may be a random unpredictable choice or may depend on whether fight scene 465 has been used recently or not.

DESCRIPTION OF THE VOICE RECOGNITION UNIT

The embodiments of the present invention shown in Figs. 1 and 2 include voice recognition unit 38 which need only distinguish a few words such as "yes" and "no" at each branch point to accomplish a two-way dialog between each viewer and the apparatus. These words may be selected from a vocabulary of thousands of words and may be different for each branch point. But the number of alternative words that can be recognized at a given branch point should be limited to only only a few phonetically distinct words, preferably less than seven, so that voice recognition unit 38 need not distinguish among all the words in the vocabulary but only those few alternative words at each branch point. Prior-art voice recognition devices such as described in U.S. Patent 3,943,295 can be used for unit 38.

To minimize cost, a more simple voice recognition device can be used which recognizes two words at each branch point. Each pair of words may be chosen so that

one word contains a distinctive phonetic feature such as an /f/ or /s/ phoneme while the other word does not. In the example shown in Fig. 10, the screen actor suggests that the viewer say either "fight" or "run" which are easily distinguished because "fight" begins with an /f/ phoneme. The word "fight" is used rather than "attack" to make it more easily distinguishable from "run". In this embodiment the word to be recognized can be segmented into two or more intervals of 100-600 milliseconds each, during which a count is made of zero voltage crossings. A zero count greater than a specified threshold for the first interval signals a code on line 37 to scheduler 35 that a fricative was used.

More elaborate word-recognition methods may be used by unit 38. For example, devices using two or more bandpass filters, fast Fourier analysis, autocorrelation, or other prior-art voice recognition methods may be used. The decision-making logic of recognition unit 38 may include decision trees, decision matrixes, best-fit template matching, and/or other methods for determining which preprogrammed combination of voice characteristics or features most resembles the sound spoken by the human viewer.

These characteristic features may include isolated words, words in continuous speech, syllables, phrases, non-word voice sounds, and/or a count of the number of phonemes or phoneme/phoneme combinations in the received sound. The presence of any sound above a given threshold may be used as a feature. If syllable recognition is used, the set of prompting words at each branch point should be planned so that each word uses a syllable or combination of syllables not found in any of the other words at that branch point.

At some branch points it may be appropriate for the viewer to speak whole phrases which may be displayed as a list of alternative prompting messages on screen 27 (Fig.

1) or display 174 (Fig. 13). Unit 38 may analyze only the first word of the phrase or may use prior-art methods of recognizing keywords in continuous speech. There is no need for unit 38 to recognize every word in the phrase, because the alternatives are restricted to only a few words or phrases at each branch point.

FUNCTIONAL DESCRIPTION OF COMMAND PROCESSING

Control information includes story commands and cue commands. Cue commands specify what is to happen during an interval of time. Story commands represent points in time, and form chains which define each alternative story line. Branch points in the movie, when a viewer can choose among alternatives, are represented by special story commands which can point to several subsequent chains of story commands. This results in a complex network of story command chains illustrated in Fig. 5.

Story commands may consist of a prefix followed by one or more addresses or data. Cue commands may be fixed or variable length records which are modified and moved to cue table 31 by scheduling unit 35. Story commands will often contain pointers to cue commands. These pointers tell scheduling unit 35: "Schedule this cue command for this point in time". The time interval represented by each cue command is relative to all that has come before it. Thus if a cue command is inserted into a chain it displaces all subsequent cue commands in time. Several cue commands may begin at the same point in time (synchronized video and audio for example). The story commands pointing to such synchronized cue commands are chained together and are stored in memory 85 one after the other in any convenient order.

In contrast to cueing unit 12 which executes the cue commands at the instant their start time arrives, scheduling unit 35 processes the story commands several seconds ahead of the start time. As scheduling unit 35 processes the story commands in each chain, it does not

immediately cause a video or audio event to happen. Rather, scheduler 35 schedules that event by determining when the cue command should cause the event to happen.

When scheduling unit 35 processes a story command, it follows the chain of pointers to various cue commands to determine which optical read head 51 or 54 should be committed to which video track during which time interval, so that any blocks of audio/graphics that are required during that interval can be scheduled in advance while the head is still available. Scheduler 35 cannot schedule video far beyond each branch point in the movie because there would be many more possible video sequences than there are heads to read it. But the control blocks and audio for every possible choice at the next branch point should be read into memory 85 and 86 in advance of the branch point so that when the viewer makes a decision, the audio for line 22 can be generated without delay. Also a read head should be moved into position in advance to cover all alternative video tracks which may be required after the branch. This advance scheduling insures that there is no discontinuity in either the video or audio and that both remain synchronized through the cue table rescheduling which scheduler 35 does after each decision by a viewer.

To illustrate how one embodiment of the apparatus may recycle video frames and synchronize them with alternative audio tracks, consider the example in Fig. 7 in which the video sequence is the talking head of an actor and the audio tracks are his (or someone's) voice. Time is represented in Fig. 7 as flowing left to right. Strip 323 represents a sequence of video frames as they are recorded on the videodisc. Rectangle 301 represents one such video frame. But the frames are not read in strip 323 sequence; rather the frames are read first in strip 322 sequence through frame 313, then in strip 324 sequence from frame 303 through frame 314. Frame 303 is

UU16514

read a second time immediately following frame 313 because five of the frames (303, 306, 307, 312, 313) are used twice, thereby saving five tracks of videodisc space. Audio block 321 is synchronized with video sequence 322 and audio block 325 is synchronized with video sequence 324. The actor's head may move about during sequence 322, but frame 303 is chosen so that the head position of frame 313 is about the same as in frame 303, so that there is no sudden movement at the electronic splice. The whole sequence of frames 301 through 314 requires seven cue commands, because there are five series of consecutive video frames to be addressed by cueing unit 12, and two audio blocks to be synchronized with the video. Unit 35 schedules the first video frame 301 and audio 321 to begin in synchronism. A fractional frame 320 of audio block 321 is automatically trimmed to synchronize audio with the video frames which begin with frame 301.

If video frames 301 through 314 were merely repeated as many times as needed to cover all the audio, something resembling badly synchronized foreign-language dubbing would result. The reason that frames 304 and 305 are skipped in sequence 322 and frames 308 and 309 skipped in sequence 324 is to best match the available inventory of video frames to each block of audio. To establish a precise match between the phonemes of each portion of digitized audio and a partial selection of video frames containing the closest approximation to the natural lip positions for those phonemes, the cue commands select video frames in the same sequence as the original video recording, but with many frames skipped. The cue commands then reuse the same frame sequence (perhaps in reverse order) with different frames skipped.

Audio also requires automatic show-time editing, especially whenever frames of audio are inserted into a continuous audio sequence. Several alternative audio

inserts may be used which require slightly different timing. Also these audio inserts may be used with many different audio tracks each of which has a slightly different speech rhythm. An insert which starts at just the right instant in one sequence may cause an undesirable lag in another sequence. To correct this problem the cue command which invokes the audio insert also specifies how many eighths of frames of audio to omit at the beginning and end of the insert. Alternative audio inserts may each have different lengths which may require lengthing or shortening of the video frame sequence to preserve lip-synchronism. Each of these audio/video combinations may be specified by one pair of cue commands.

Each cue command in the illustrated embodiment is a fixed-length record of binary coded data and represents an interval of time that is scheduled to begin at the instant indicated within the cue command. There is at least one cue command for each series of consecutive video frames and for each portion of audio. One scene may require hundreds of commands which are selected and stored into cue table 31 by scheduling unit 35 and executed by cueing unit 12. Cue table 31 is therefore similar to a first-in/first-out queue, except at branch points in the movie when a viewer's decision may cause scheduling unit 35 to abandon several commands in cue table 31 (representing video and audio not yet presented) and to replace them with several new commands representing the altered story line.

DETAILED DESCRIPTION OF THE SCHEDULING UNIT

The detailed structure of one embodiment of scheduling unit 35 is shown in Fig. 3. Scheduling unit 35 receives blocks of digitally coded control data from retrieval unit 55 which stores story commands into random-access memory (RAM) 85 via line 83 and stores cue commands into memory 86 via line 84. Memories 85 and 86

are shown in Fig. 1 as a single box. Memory 86 may be an extention of memory 85 but the two memories are distinguished in Fig. 3 for clarity.

The course of a movie is controlled by structures of story commands in memory 85. There are at least two kinds of story commands: commands which represent branch points in the movie, and pointers which point to cue commands and other story commands. Each kind of story command is read from memory 85 at a location specified by counter 82 which is incremented via line 72 by control circuit 62 so that chains of story commands in memory 85 are sequentially addressed for processing in register 65 or 78. Registers 65 and 78 may be conventional random access memory (RAM) working storage, but are shown separately in Fig. 3 for clarity.

A story command addressed by counter 82 is moved from memory 85 via bus 74 to register 65. The left-most byte (herein called the "prefix") of the story command in register 65 is moved via line 63 to control circuit 62 (to command decoder 530 in Fig. 6) which distinguishes branch commands from pointers. If the prefix on line 63 indicates a pointer, the story command is moved from memory 85 via bus 80 to register 78. The left pointer address of the story command in register 78 specifies a location of a cue command in memory 86. This cue command is addressed via line 79 and is moved via line 87 to register 90 for insertion of the start time (which will appear on line 105 in Fig. 4). The right pointer address of register 78 specifies the next story command in the chain of pointers (illustrated in Fig. 5).

Each cue command represents an interval of time which is relative to the intervals which have preceeded it. The sum of all these prior intervals is the time at which the next interval will be scheduled. This cumulative time is stored in register 91 in units of 1/30 second. When a new cue command is moved to register 90,

the start-time field 88 is initialized via line 92 with the cumulative time value in register 91. Register 91 is then updated by adder 94 which adds the duration field 89 from register 90 to register 91 via lines 95 and 93. Register 91 now represents the point in time immediately following the time interval for the cue command in register 90. This cue command is moved from register 90 via line 32 to cue table 31 at the next available location indicated by counter 97 which addresses cue table 31 via line 98. Control circuit 62 then increments counter 97 via line 64 to the next available unused location or to the location of an old completed cue command whose space in cue table 31 may be reused. Control circuit 62 also increments counter 82 via line 72 to address the next story command in memory 85. When the end of the block of story commands in memory 85 is reached, control circuit 62 updates track address register 47 via line 48 and requests the next block of commands from retrieval unit 55 specified to tracking unit 58 by the track address on line 49.

Each cue command may be located in memory 85 immediately following story command prefix 96 to avoid need for unnecessary pointers. This arrangement is used in Fig. 5. But in Fig. 3 the cue commands are explicitly pointed to by the left pointer in register 78 and are assigned separate memory (block 86) from the story commands (block 85) to clearly distinguish story command processing from cue command processing. The right pointer of the story command in register 78 specifies a successor story command in a chain of story commands. The right pointer in register 78 is moved via lines 75 and 73 to counter 82 which addresses via line 81 the successor story command in memory 85.

Referring to Fig. 5, a schematic flow diagram is shown for a typical chain or network of story commands. In contrast to the apparatus blocks in Figs. 1

through 4, the blocks shown in Fig. 5 represent data, specifically story commands, and the arrows represent associative relationships between the commands. Blocks 200, 202, 203, etc. are pointer story commands which in Fig. 3 are sequentially read from memory 85 and processed in register 78. Blocks 204 are branch story commands which in Fig. 3 are processed in register 65. The various command prefixes shown in Fig. 5, such as prefix 96, indicate what kind of story command it is. The prefixes are abbreviated herein as B for Branch, W for Wait, D for Do, C for Cue, and E for End.

The branching chain shown in Fig. 5 consists of a horizontal chain of right pointers, and vertical chains of left pointers. At the end of each branch of each chain is one or more cue commands, such as video cue commands 214 and audio cue commands 217 and 220. At the end of the last episode of the movie there may be a final schedule of pointers which does not branch, but instead shuts down the system.

The reason the branched chain shown in Fig. 5 is arranged in columns linked together horizontally is to emphasize an important distinction. Some events must happen sequentially (such as sequences of video frames), but other events must happen concurrently (such as synchronized audio and video). The horizontal chain at the top of Fig. 5 (blocks 200 through 206) represent events to be scheduled for sequential execution by cueing unit 12. Each vertical chain in Fig. 5 (blocks 210 through 219) represent events to be scheduled for concurrent execution by cueing unit 12. At the end of each branch there are one or more (usually several) cue commands (such as block 214) which are executed sequentially. At the end of each such sequence there is a one-byte E prefix (215, 218 and 221 in Fig. 5) which is passed via line 63 in Fig. 3 to control circuit 62 instructing it to discontinue the sequence of cue

commands addressed via line 79 by the left pointer in register 78, and instead, instructs control circuit 62 to proceed to the next column in the chain specified by the right pointer in register 78 which via lines 75, 73 and 81 addresses the next story command in memory 85. For example, in Fig. 5 when all of the E prefixes (215, 218 and 221) have been reached in the scheduling of the commands in the column headed by story command 200, command 200 is returned to register 78 in Fig. 3 to obtain the right pointer (line 201 in Fig. 5) which addresses the next story command 202 from memory 85. Command 202 replaces command 200 in register 78, and processing continues with the second column in Fig. 5 (headed by block 202).

Since story commands 210, 216 and 219, which are chained together via their right pointers, each contains a D prefix (for Do), each of the chains of cue commands pointed to by their left pointers is scheduled to begin at the same point in time (specified by register 91 in Fig. 3). Typically, the video frames pointed to by cue commands 214 will be sequentially displayed, but this video will run concurrently with the audio pointed to by cue commands 217, and also concurrently with the audio pointed to by cue command 220. Command 220 may point to the digitized name of one of the viewers as spoken by the same actor whose digitized voice is pointed to by commands 217. Command 220, in other words, may represent an audio insert. The video frames pointed to by commands 214 are preselected to best synchronize with audio 217 and 220 for consistency with lip movements, facial expressions, gestures, tone of voice, etc.

The W prefix (for Wait) in story command 200 instructs control unit 62 not to read command 202 into registers 65 and 78 until after all the E prefixes 215, 218, 221 subordinate to command 200 have been reached. The right pointer of the last story command in each

001631

vertical chain (such as 219) has an X in it, which is a null pointer indicating the end of the chain.

Story commands 204 in Fig. 5 represent a branch point in the story line which can lead to several alternative chains of story commands (such as 206 and 208) depending on the viewer's choices. Referring again to Fig. 3, when a story command is moved from memory 85 via bus 74 to register 65 the prefix is moved via line 63 from register 65 to control circuit 62 (to decoder 530 in Fig. 6). Several types of branch commands may be used. The branch code prefix on line 63 may indicate an unconditional jump, in which case the memory address in counter 82 is replaced via lines 67 and 73 with the branch-address field from register 65.

Most branch commands will represent decision points in the movie when the viewer can input a verbal response through microphone 40 (Fig. 1) or through push buttons 42. These signals are represented in Fig. 3 on lines 37 and 44 respectively as a 4-bit binary code which is passed via line 71 to comparator 69 which compares the binary code on line 71 with the condition code on line 68 from a succession of branch commands in register 65. If an inappropriate response code is present on line 71 it will not match any codes on line 68 and will therefore be ignored. If no new response is entered by the viewer, control circuit 62 will not receive the response code via line 70. Control circuit 62 decrements timer 60 which imposes a time limit (of a few seconds) on the viewer's response, i.e. while RS flip-flop 532 in Fig. 6 is set. During this period a true signal on line 531 inhibits sequential cycle controller 533 from proceeding to the next series of pointer commands so that the branch commands recycle through register 65. This loop is indicated by boxes 420 and 421 in Fig. 9. When the time limit expires in timer 60, control circuit 62 forces a default response code onto line 71 via lines 161 and 59

so that comparator 69 will detect a match with one of the branch commands in register 65.

When comparator 69 detects a match, it enables gate 77 via line 76 which causes the branch address field in register 65 to replace the address in counter 82 via lines 67 and 73. The next story command obtained from memory 85 at a location specified by the new branch address in counter 82 and bus 81, will be a new series of pointer commands for register 78 which represent the new story line appropriate to the viewer's response or lack of response.

Story commands which test previously-set conditions may be used to permit a variable number of viewers. Each viewer plays the role of a character whose pre-recorded voice and images are bypassed if a human is playing that role. After the viewers inform the microcomputer in Fig. 1 (through a series of questions and answers) of how many viewers there are and who is playing which role, this information can be tested frequently using branch commands which cause branch address 67 to be taken if a human viewer is playing that role, but proceed to the next sequential branch command if the role is to be played by a prerecorded image of an actor(s).

DESCRIPTION OF THE SCHEDULING UNIT PROGRAM

In the preceeding section the structure of scheduling unit 35 was described using separate components (Fig. 3). An alternative embodiment is a programmed microprocessor which performs processing equivalent to that described in the preceeding section by performing a sequence of steps such as the program sequence shown by flowchart Fig. 9.

Referring to Fig. 9, the story commands and cue commends are read into memory during step 401. These may be read together when power is first turned on, or may be read piecemeal. Step 402 tests the prefix of the first story command for the code "B" or a numeric equivalent.

When the loop indicated by line 412 encounters a B (Branch) command, control proceeds to step 413 (described below). Otherwise, control proceeds to step 403 which stores the W (Wait) command into working storage. Block 200 in Figure 5 represents the kind of W command being processed at this point. The left address of the W command is a pointer to the D command (block 210 in Fig. 5). This D command is picked up in step 404. Step 405 then checks whether the left address of the D command points to a cue command (block 212 in Fig. 5). If it does so, control proceeds to step 406 which schedules the cue command by storing it in cue table 31 (see Fig. 1) after modifying the start time field 88 as described in the preceeding section. After step 406 has finished, step 407 checks the next cue command which immediately follows command 212 in memory 85. If the prefix is not an E (for End), control loops back to step 406 to schedule another cue command. If it is an E, control proceeds to step 408 which checks the right address of the D command got during step 404. If the right address points to the next D command (block 216 pointed to by address 213 in Fig. 5), control loops back to step 404 (via line 409) to get the next D command.

Steps 404 through 408 continue to loop in this sequence until a D command is encountered which does not have a pointer in its right address (block 219 in Fig. 5). When step 408 encounters such a D command it passes control to step 410 which restores the W command saved by step 403. The right address of this W command is used in step 411 as a pointer to the next W or B command (block 202 pointed to by address 201 in Fig. 5). But if it is an E code step 426 terminates the show by passing control to step 427 which stops the apparatus. Otherwise, control loops back to step 402 which checks the new story command picked up by step 411.

If this command is a B command like block 204 in

Fig. 5, step 402 then passes control to step 413 which checks whether the audio blocks pointed to by cue commands for the current B command has been read by retrieval unit 55 into memory 125. If not, step 414 requests retrieval unit 55 to read this block of audio. Step 415 then checks whether electro-optical read heads 51 and 54 have been positioned for the video frames if the current B command will match the choice code sent to scheduling unit 35 from hand-held input device 41. At this stage in the processing (before the viewer has made a choice) all contingencies should be prescheduled in cue table 31. If step 415 finds that the read head is not yet scheduled, step 416 is performed which stores a head-positioning cue command into cue table 31. Step 417 then saves the B command in working storage for later use by step 411. The next byte of memory after the B command is then checked in step 418 for an E (end) code. If another B command is found, control loops back (line 419) to step 413 to process the next B command. Steps 413 through 418 continue to loop through several B commands until the E code is encountered, at which time control is passed to step 420. Step 420 checks signal bus 70 in Fig. 3 for an indication that the viewer has made a choice. If he has, control proceeds (via line 425) to step 424. If no choice has occurred, timer 60 is checked in step 421. If the time limit has elapsed, control proceeds to step 424. Otherwise control loops back (via line 422) to step 420. Loop 422 continues until either the time elapses or the viewer makes a choice. In either case, step 424 searches the B commands saved during step 417 for a match with the choice code on bus 70 in Fig. 3. If no match is found, the viewer is incorrectly making a choice which is not used at this branch point so the choice is ignored by continuing the 422 loop. When a choice is found by step 424 (which may be the default choice forced by step 421, control proceeds to step 411

which picks up the address of the next W command (block 208 pointed to by address 207 in Fig. 5).

DETAILED DESCRIPTION OF THE CUEING UNIT

The detailed structure of one embodiment of cueing unit 12 is shown in Fig. 4. A program flowchart for cue command processing is shown in Fig. 14 which illustrates one of many sequences of steps which may be used to perform the functions of cueing unit 12. Referring to Fig. 4, each cue command is moved one at a time from cue table 31 via line 103 into buffer 102 which may consist of several bytes of conventional random-access memory (RAM) or a special purpose register. The bits of buffer 102 are arranged in fields of one or more bits which are processed via lines 104-115 in Fig. 4.

At the end of each video frame, circuit 10 sends a signal via line 140 to increment real-time frame counter 138, a conventional binary counter. This signal may be generated at the end of each field if desired. The time value in counter 138 is compared in comparator 136 to the start time bits on line 105 from buffer 102 for each cue command. If comparator 136 determines that the start time value on line 105 is greater than or equal to the real-time value on line 137 it sends an initiating signal via line 160 to initiator switching circuit 131. This initiation signal is suppressed if the 3-bit status code on line 104 indicates that the command is to be ignored. Conversely if the status line 104 indicates that the command is to be executed immediately, comparator 136 sends an immediate initiating signal via line 160 to initiator 131.

Initiator circuit 131 (detailed in Fig. 12) generates various switching signals on lines 132-135 depending on the 3-bit start code on line 106, the 3-bit command-type code on line 107 and the 1-bit channel indicator on line 108. If the type code 107 indicates a video command, initiator 131 leaves audio memory control

line 132 unchanged, enables video fader control lines 133, enables video switch 142 via lines 134, and selects the read head indicated by channel bit 108 via line 135 and tracking unit 58. If start code 106 indicates "take" and channel bit 108 indicates a change in read heads, initiator 131 signals video switch 142 via lines 134 to switch off the video channel for the frame just completed and to switch on the other channel after tracking unit 58 has positioned the read head to the track addressed on bus 139. This 18-bit track address is obtained via bus 109 from the command in buffer 102. If no head change is specified by channel bit 108, the video switch signal on lines 134 remains unchanged. If start signal 106 indicates "fade in", initiator 131 signals video fader 148 via lines 133 to gradually increase the amplitude of the composite video signal on line 141. If start signal 106 indicates "mix", the control signals on lines 133 and 134 remain unchanged for the occupied channel, but signal the switch 142 or fader 148 to perform a take or fade-in for the newly selected channel specified by channel bit 108. A mix without a head change implies there is only one read head in this unit, so initiator 131 will treat this command as a take. If start code 106 indicates "cancel", the switch 142 and fader 148 are controlled by terminator circuit 118 which is described below.

The chroma invert signal on line 110 changes whenever an odd number of video frames are skipped, to avoid loss of chroma phase lock. Signal 110 causes conventional inverter circuit 145 to shift by 180 the phase of the chroma portion of the composite video signal on line 11, and recombine the inverted chroma with the luminance portion, so that the composite signal on line 147 will continue in phase with the color subcarrier. The invert signal on line 144 causes video circuit 10 to invert the burst signal to be in phase with the subcarrier.

If type signal 107 indicates a head-positioning command, control lines 133 and 134 remain unchanged for the occupied channel, so the video signal passing through blocks 58, 10, 145, 148, and 142 will not be disturbed. Head selection signal 108 is passed via initiator 131 and line 135 to tracking unit 58 in conjunction with the track address on buses 109 and 139. Tracking circuit 58 then positions the unoccupied read head to the track address specified on bus 139. However, switch 142 for the selected channel is not enabled.

If type signal 107 indicates "go to", the cue command in cue table 31 located at the relative address given on bus 109 is loaded into buffer 102 via line 103 replacing the current "go to" command and is given "immediate" status on line 104. The "go to" command is given "defer" status in cue table 31 via line 101 by terminator 118.

If type code 107 indicates an audio or graphics command, initiator 131 leaves lines 133-135 unchanged for both video channels, and enables audio/graphics memory 125 via control line 132. Address 109 which is used as a disc track address for video commands has a different use for audio commands. Address 109 indicates the location of data blocks in memory 125 which is a conventional random access memory (RAM) into which blocks of digitally coded compressed audio and graphics data are stored via line 56 by retrieval unit 55 which obtains this data from non-picture tracks on disc 52. Plug-in cartridge 15 which contains conventional non-volatile RAM is addressed via bus 130 as an extention of memory 125. The RAM in cartridge 15 contains digitized audio recordings of the viewer's names as spoken by the various actors that may use the viewer's names during the show. Line 16 indicates that cartridge 15 is an extension of memory 125. When memory 125 is read-enabled by initiator 131 via line 132, memory 125 treats the binary address on bus

130 as a memory address of the first byte of data which it outputs on bus 128 or 127 (shown separately for clarity) depending on whether the data is audio or graphics. If audio, the byte on bus 128 is used by audio mixer 129 which increments the address on bus 130 via line 161 to access as many bytes from memory 125 as are needed to form a continuous audio signal. Mixer 129 edits the data from memory 125 according to the values on lines 111-113 from the command in buffer 102.

The "trim" field 111 in buffer 102 indicates the amount of audio signal to be trimmed from the beginning of the audio recording by mixer 129 to achieve precise synchronization for the current combination of audio and video. Mixer 129 performs this trimming while the digitized audio is still in compressed form in memory 125. Although each block of digitized audio in memory 125 begins at a frame boundary, i.e. at 1/30 second intervals, this resolution is too coarse for precise audio editing, especially where variable-length spoken words must be inserted into dialog. The trim value on line 111 therefore represents eighths of video frames. Mixer 129 discards the amount of audio indicated by trim field 111 and stores the trimmed series of bytes into memory 120 via line 119. Memory 120 may be a continuation of conventional RAM 125 but is shown separately in Fig. 4 for clarity. Mixer 129 may also attenuate the digitized audio by multiplying each byte by the attenuation factor on line 112 from buffer 102.

After mixer 129 stores the edited digitized audio bytes into memory 120 and subsequent commands perhaps have caused mixer 129 to add additional audio to the bytes in memory 120, circuit 124 moves blocks of audio data from memory 120 into fast-in slow-out register 19. Register 19 may be a conventional charge coupled device (CCD) which is filled with digitized audio via line 18 at a bit rate of about 12 megHz and readout at a sampling

(byte) frequency of about 10-12 kiloHz. Register 19 may also be a conventional RAM which is readout at the 10-12 kiloHz rate.

If type code 107 indicates a graphics command, memory 125 passes a series of bytes via line 127 to graphics generator 126, which may be a conventional circuit such as used with video games and which generates video signals on line 146 corresponding to various shapes, alpha/numeric characters and lines for display on TV screen 27. The binary data on line 127 may consist of raster coordinates, color selection, selected character/shape, color of character, orientation, brightness, direction of motion, and speed. For embodiments in which animated cartoons substitute for camera-originated frames, graphics generator 126 generates video frames containing cartoon images.

Duration field 113/114 consists of a binary number which specifies the length of the time interval the current command is to be active. This number represents frame counts in video commands and eighths of frames for audio commands. For video commands counter 117 is initialized via line 114 with the duration count from buffer 102. The end-of-frame signal on line 100 decrements counter 117 each 1/30 second. When zero is reached, counter 117 signals terminator switching unit 118 via line 116 to begin the termination sequence. For audio/graphics commands the duration field in buffer 102 is moved via line 113 to mixer 129. When mixer 129 has counted down the duration value from line 113, mixer 129 signals terminator 118 via line 162.

When terminator 118 (detailed in Fig. 12) receives a signal on line 116 it begins the termination sequence specified by the finish code on line 115. Terminator 118 also voids the status code of the current command in cue table 31 via line 101 so that cueing unit 12 will not move the completed command again from cue table 31 and to

indicate to scheduling unit 35 that the cue table space may be reused.

If type signal 107 indicates a video command and 3-bit finish code 115 indicates "cut", terminator 118 signals video switch 142 via line 121 to switch off video signal 144 for the channel indicated on line 108. If finish code 115 indicates "fade out", terminator 118 signals fader 148 via line 122 to gradually reduce the amplitude of the video signal on line 141 and then switch it off in circuit 142 (after delay 540 in Fig. 12, of 2 seconds or so). If finish code 115 indicates "repeat", lines 121 and 122 remain unchanged, but the track address on bus 139 is reinitialized to the buffer 102 value on bus 109, and counter 117 is reinitialized with the duration value on line 114. Thus the video frame sequence (or freeze frame if the duration is one) is restarted from the initial frame, except that the start signal on line 106 is not reprocessed. If finish code 115 indicates "next", the next sequential cue command in cue table 31 is loaded into buffer 102 via line 103 and given "immediate" status on line 104. The status of the command just terminated is set in cue table 31 by terminator 118 via line 101 to a "defer" status.

If type code 107 indicates an audio/graphics command, video control lines 121 and 122 remain unchanged. A "cut" signal on line 115 tells mixer 129 to stop editing digitized audio from memory 125. A "fade out" tells mixer 129 to gradually attenuate the edited audio in memory 120 just as if the attenuation value 112 were decreasing. "Repeat" and "next" are processed the same for audio/graphics commands as for video.

DESCRIPTION OF THE CARTOON GRAPHICS GENERATOR

Referring to Fig. 2, an alternative embodiment is shown in which the branching movie is an animated cartoon digitally generated by graphics generator 126 from compressed digitized data which may be read along with

digitized audio from videodisc 52 and/or from other mass-storage devices such as magnetic bubble memory 173.

Cueing unit 12 executes the cue commands at the times specified therein by conveying to cartoon generator 126 via line 127, blocks of compressed binary-coded data previously stored into memory 125 by retrieval unit 55 and used by generator 126 to generate one or more cartoon frames.

Circuitry for reading standard video, blanking, burst and sync from disc 52 is not required in this embodiment because the video signal is generated on line 146 (Fig. 4) by generator 126. The information read by conventional tracking unit 58 and retrieval unit 55 may consist entirely of compressed digitized data from which video, audio, prompting messages, and other signals are generated.

The graphics generator chip manufactured by General Instrument Corp. for their GIMINI video games is suitable for unit 126 in simple embodiments of the present invention.

The data compression method used for storing animated cartoon data, is a line-by-line string coding method in which much of the redundancy in each raster line is removed. A similar coding method is described in "Raster Scan Approaches to Computer Graphics" by Nicholas Negroponte, Computers and Graphics, Vol 2, pp 179-193, 1977. Many data compression techniques known to the art may be used in lieu of string coding. For example a catalog of 2-dimentional dot matrices may be used as described in U.S. Patent 4,103,287. Each dot matrix may include lines, corners, color background, etc. from which each cartoon frame is constructed.

DESCRIPTION OF THE CARTRIDGE

Plug-in cartridge 15, shown in Fig. 4 as an extension of memory 125, may be a non-volatile memory housed in a protective plastic case and used for storing

digitized audio recordings of the names of the various viewers as spoken by the various actors in the movie. A speech synthesis unit may be used to convert the digitized names in cartridge 15 to an audio signal on line 20.

Although it would be possible to store an entire catalog of common names and nicknames on videodisc 52 purchased by the viewers, for economy reasons the catalog of names may be stored on a second videodisc which is used by each retailer with a customizing micro-computer. The retail clerk gets a list of the viewer's names from the customer. The clerk keys these names into a keyboard connected to the customizing computer which reads the selected recordings from the retailer's videodisc and stores them into the cartridge. The customer buys this customized cartridge 15 with videodisc 52.

Digitized voice recordings of each viewer's voice may also be stored into the cartridge by the retailer's computer. The words and their phoneme components may be strung together later by cueing unit 12 to form sentences in each viewer's voice whenever the viewer pushes a button 42 to ask a prerecorded question or to make a prerecorded remark. Scheduler 35 selects cue commands which point to the digitized recordings of a viewer's voice in cartridge 15 memory depending on which button 42 is pressed in which hand-held unit 41.

Accompanying each block of digitized audio in cartridge 15 may be several schedules of cue commands which identify the video frame sequence which synchronizes with each instance of the spoken name in the digitized audio. Each instance of a viewer's name may require different video frames and hence a separate schedule of cue commands.

Although I have described the preferred embodiments of my invention with a certain degree of particularity, it is understood that the present disclosure has been made only by way of example and that equivalent embodiments and numerous changes in the details of the design and the arrangement of components may be made without departing from the spirit and the scope of my invention.

CLAIMS

1. A method of simulating a voice conversation between a previously recorded sound movie and a human viewer of the movie, comprising the steps of:

presenting to said viewer a previously recorded first scene in said movie linked to a plurality of previously recorded second scenes therein, the first scene including a face and voice speaking words to elicit from the viewer a spoken response corresponding to one second scene in said plurality thereof;

analyzing said spoken response to determine which second scene in said plurality thereof corresponds to said spoken response; and

presenting to the viewer said corresponding second scene including said face and voice speaking words responsive to said spoken response, thereby simulating a voice conversation between the viewer and the movie.

2. A method of simulating a voice conversation between a previously recorded sound movie and a human viewer of the movie, comprising the steps of:

presenting to said viewer a first scene in said movie linked to a plurality of second scenes therein;

presenting to said viewer a plurality of messages, each message corresponding to one second scene in said plurality thereof and each second scene including spoken words responsive to the corresponding message, said messages eliciting from said viewer a spoken response;

analyzing said spoken response to determine which selected message in said plurality of messages includes a word which resembles a portion of said spoken response; and

presenting to said viewer a second scene in said plurality thereof which corresponds to said selected message, the second scene including spoken words responsive to said selected message, thereby simulating a voice conversation between the viewer and the movie.

3. A method of simulating a voice conversation between a previously recorded sound movie and a human viewer of the movie, comprising the steps of:

presenting to said viewer a first scene in said movie linked to a plurality of second scenes therein;

presenting to said viewer a plurality of messages, each message corresponding to one second scene in said plurality thereof and each second scene including spoken words responsive to the corresponding message;

receiving from said viewer a response signal corresponding to a selected message in said plurality of messages;

presenting to said viewer a voice sound including words in said selected message; and

presenting to said viewer a second scene in said plurality thereof which corresponds to said selected message and which includes spoken words responsive to said selected message and said voice sound, thereby simulating a voice conversation between the viewer and the movie.

4. A method of simulating a voice conversation between a previously recorded sound movie and a human viewer of the movie, comprising the steps of:

presenting to said viewer a first scene in said movie linked to a plurality of alternative video portions and audio portions of a second scene in said movie;

presenting to said viewer a plurality of alternative verbal signals, each alternative verbal signal corresponding to one video portion and one audio portion in said plurality thereof including voice sounds responsive to the corresponding verbal signal;

receiving from said viewer a response signal corresponding to a selected verbal signal in said plurality thereof;

following the receiving of said response signal, scheduling for a point in time a video portion of said

second scene which corresponds to said selected verbal signal;

following the receiving of said response signal, scheduling for a point in time an audio portion of said second scene which corresponds to said selected verbal signal; and

presenting to said viewer at the respective scheduled points in time said scheduled video portion and said scheduled audio portion of said second scene including voice sounds responsive to said selected verbal signal, thereby simulating a conversation between the viewer and the movie.

5. An apparatus for simulating a voice conversation between a human operator of the apparatus and a previously recorded sound movie, the apparatus comprising:

means for controlling presentation of a first portion of said sound movie which is linked to a plurality of second portions thereof, the first portion including voice sounds to elicit from an operator of the apparatus a spoken response corresponding to one second portion in said plurality of second portions thereof; and

means for analyzing said spoken response and determining therefrom which second portion of said sound movie corresponds to said spoken response,

said controlling means further controlling presentation of the second portion of said sound movie which corresponds to said spoken response and which includes voice sounds responsive to the spoken response, thereby simulating a voice conversation between the movie and the operator.

6. The apparatus of claim 5, further comprising:

means for displaying a plurality of alternative responses for speaking by said operator and which correspond to said second portions of the sound movie, thereby prompting the operator to make a spoken response

which said analyzing means distinguishes from other alternative responses in said displayed plurality.

7. An apparatus for simulating a voice conversation between a human operator of the apparatus and a sound movie, the apparatus comprising:

means for controlling presentation of scenes in said sound movie, a first scene therein being linked to a plurality of second scenes therein, the first scene including voice sounds to elicit from the operator a spoken response corresponding to one second scene in said plurality thereof;

means for analyzing said spoken response and determining therefrom which second scene in said sound movie corresponds to said spoken response, and

means for scheduling for a point in time a second scene in said sound movie which corresponds to said spoken response and which includes voice sounds responsive to said spoken response,

said controlling means further controlling presentation of said scheduled second scene at said point in time, thereby simulating a voice conversation between the movie and the operator.

8. The apparatus of claim 7, wherein each scene in said sound movie comprises video portions and audio portions, the apparatus further comprising means for selecting from a plurality of alternative audio portions one selected audio portion which includes said voice sounds responsive to the operator's spoken response,

said scheduling means further scheduling said selected audio portion for a point in time in synchronism with one video portion, thereby synchronizing said one video portion with one of said plurality of alternative audio portions depending on the operator's spoken response.

9. The apparatus of Claim 8, further comprising:

means for scheduling presentation of said video

portions in accordance with a branching data structure of digital pointers which specify alternative sequences of said video portions, and wherein said analyzing means includes means for selecting a branch in said data structure in accordance with a digital pointer corresponding to said spoken response.

10. The apparatus of claim 7, wherein each scene in said sound movie comprises video portions and audio portions, and wherein the audio portions include a plurality of alternative spoken names, the apparatus further comprising:

means for selecting from a plurality of alternative audio portions one selected audio portion which includes the name of said operator,

said scheduling means further scheduling said selected audio portion for a point in time in synchronism with one video portion, thereby synchronizing said one video portion with said operator's name.

11. An apparatus for simulating a voice conversation between a human operator of the apparatus and a sound movie, the apparatus comprising:

means for controlling presentation of scenes in said sound movie, a first scene therein being linked to a plurality of second scenes therein, the first scene including voice sounds to elicit from the operator a response corresponding to one second scene in said plurality thereof;

means for displaying a plurality of alternative verbal responses which correspond to said second scenes of the sound movie;

means for receiving from the operator a response signal which corresponds to a selected verbal response in said displayed plurality thereof; and

means for generating a voice signal including words displayed in said selected verbal response, thereby simulating the operator's side of a voice conversation,

said controlling means further controlling presentation of the second scene of said sound movie which corresponds to said selected response and which includes voice sounds responsive to the selected response, thereby simulating a voice conversation between the movie and the operator.

12. An apparatus for simulating a voice conversation between a human operator of the apparatus and a sound movie, the apparatus comprising:

means for controlling presentation of scenes in said sound movie, a first scene therein being linked to a plurality of second scenes therein, the first scene including voice sounds to elicit from the operator a response corresponding to one second scene in said plurality thereof;

means for communicating to the operator a plurality of alternative verbal responses which correspond to said second scenes of the sound movie;

means for receiving from the operator a response signal which corresponds to a selected verbal response in said communicated plurality thereof; and

means for scheduling for a point in time a second scene in said sound movie which corresponds to said response signal and which includes voice sounds responsive to said selected verbal response;

said controlling means further controlling presentation of said scheduled second scene at said point in time, thereby simulating a voice conversation between the movie and the operator.

13. An apparatus for simulating a voice conversation between a human operator of the apparatus and a sound movie, the apparatus comprising:

means for generating an audio signal of a voice speaking a plurality of words to elicit from a human operator a spoken response;

means for processing a video signal for presentation

with said audio signal as a sound movie which includes an image of a speaking person;

means for analyzing said spoken response to determine which one word in said plurality of words most closely resembles a portion of said spoken response; and

means for selecting from a plurality of recorded voice sounds a selected voice sound corresponding to said one word,

said generating means further generating an audio signal which includes said selected voice sound for presentation with said sound movie.

14. An apparatus for simulating a voice conversation between a human operator of the apparatus and an animated cartoon sound movie, the apparatus comprising:

means for generating an audio signal including voice sounds which communicate to the operator of the apparatus a plurality of alternative words to speak in response;

means for generating a video signal including animated cartoon images of a talking face, wherein said voice sounds and said talking face comprise scenes in said cartoon movie;

means for controlling presentation of a first scene in said cartoon movie which is linked to a plurality of second scenes therein, each second scene corresponding to one word in a plurality of alternative words included in said first scene; and

means for analyzing a spoken response from said operator and determining therefrom which selected word in said first scene corresponds to said spoken response,

said controlling means further controlling presentation of the second scene in said cartoon movie which corresponds to said selected word and which includes voice sounds responsive to the selected word, thereby simulating a voice conversation between the cartoon movie and the operator.

15. A method of speech recognition, comprising the steps of:

presenting to a human speaker a plurality of alternative words, each word including a distinctive phonetic feature, to elicit from the human speaker a spoken response which includes one word in said plurality thereof; and

analyzing said spoken response to determine therefrom which word in said plurality of alternative words includes a distinctive phonetic feature also included in said spoken response.

16. A method of speech recognition, comprising the steps of:

presenting to a human speaker a plurality of alternative words, each word including a combination of distinctive phonetic features, to elicit from the human speaker a spoken response which includes one word in said plurality thereof; and

analyzing said spoken response to determine therefrom which word in said plurality of alternative words includes a combination of distinctive phonetic features also included in said spoken response.

0116314

1/15

Fig. 1

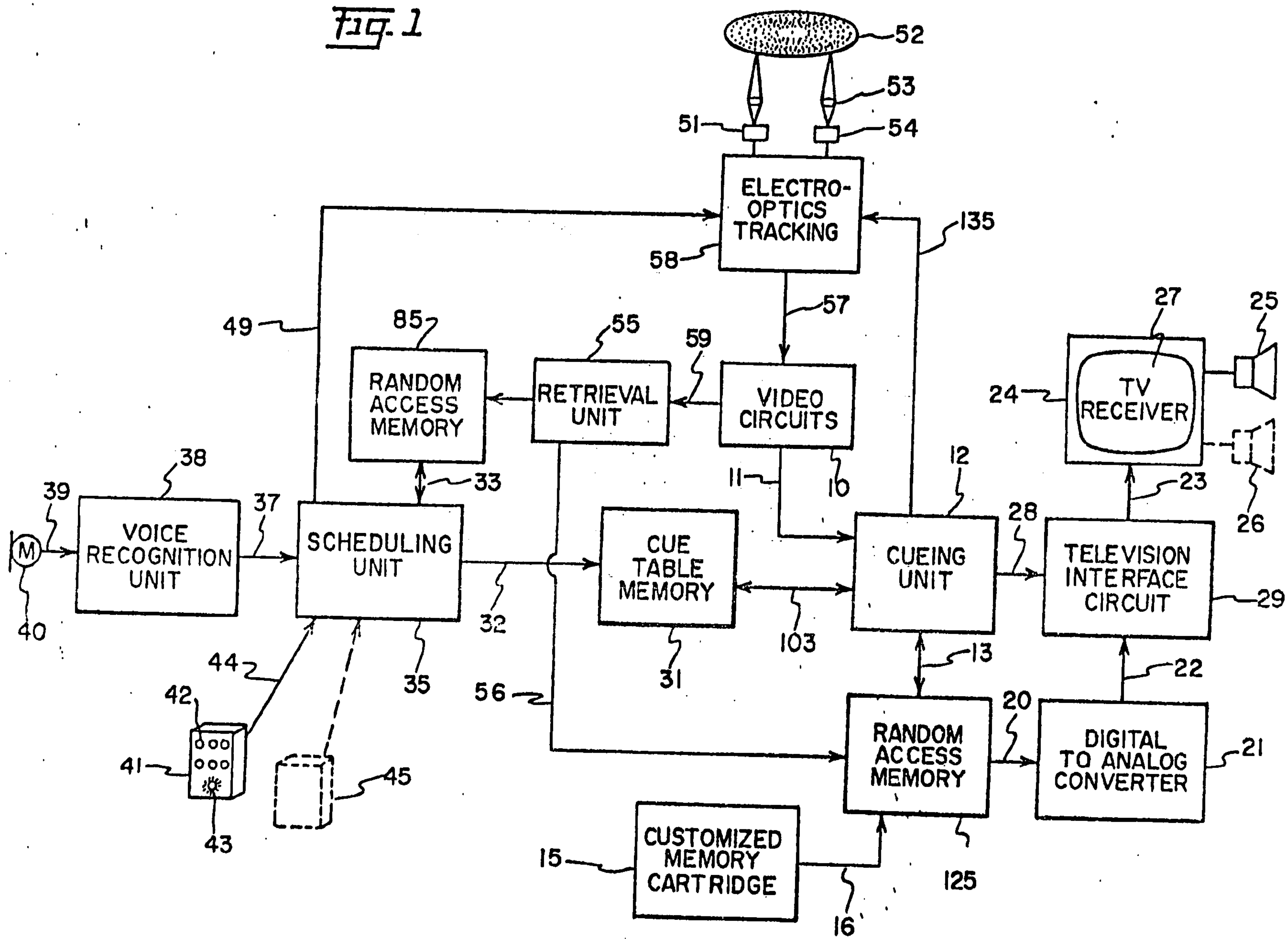
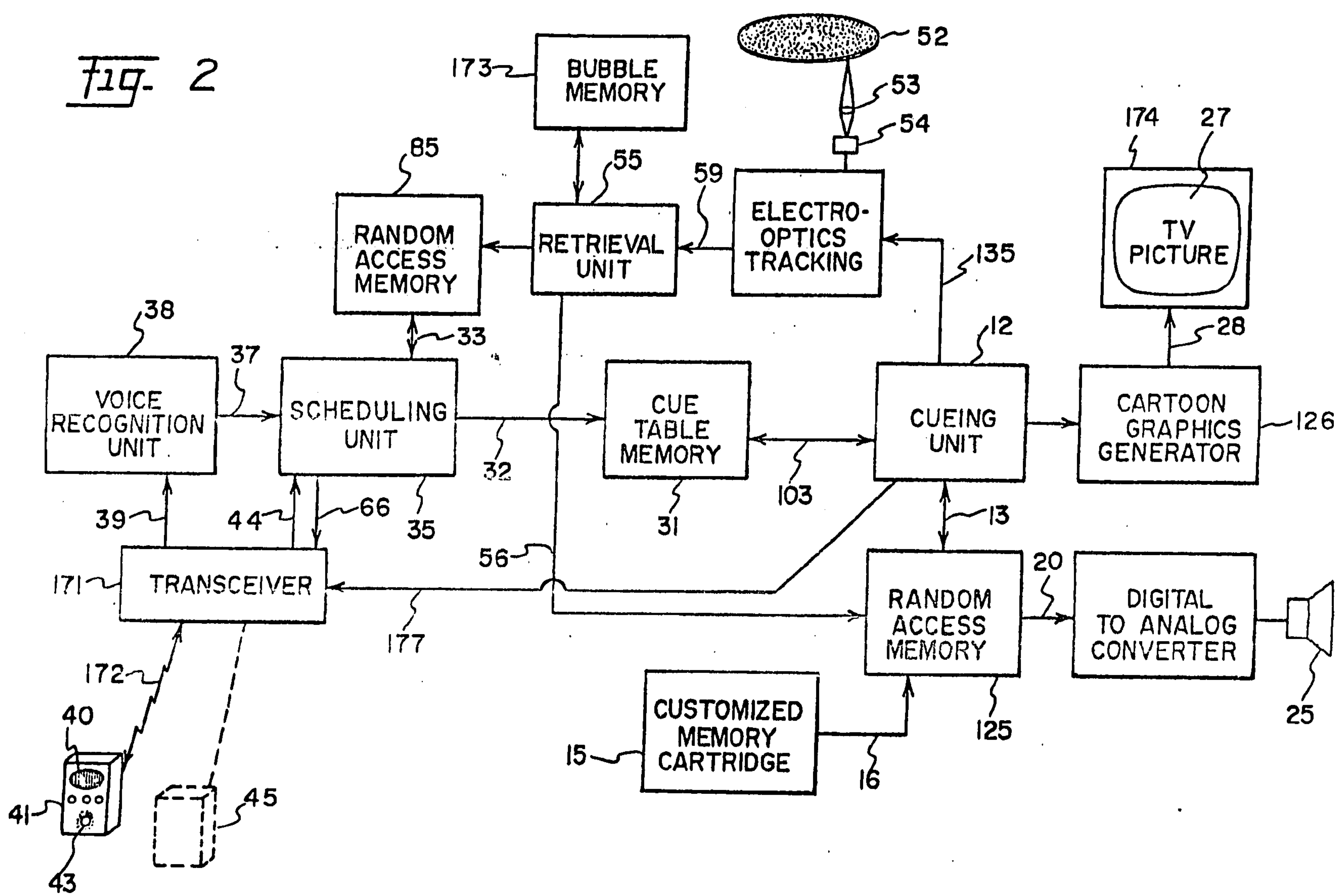


Fig. 2


BUBBLE MEMORY
173
52
53
54
RANDOM ACCESS MEMORY
85
RETRIEVAL UNIT
55
59
ELECTRO-OPTICS TRACKING
135
TV PICTURE
174
27
28
VOICE RECOGNITION UNIT
38
37
SCHEDULING UNIT
33
32
35
CUE TABLE MEMORY
31
103
CUEING UNIT
12
CARTOON GRAPHICS GENERATOR
126
39
44
66
13
TRANSCEIVER
171
56
177
RANDOM ACCESS MEMORY
125
20
DIGITAL TO ANALOG CONVERTER
25
172
40
41
43
45
CUSTOMIZED MEMORY CARTRIDGE
15
16

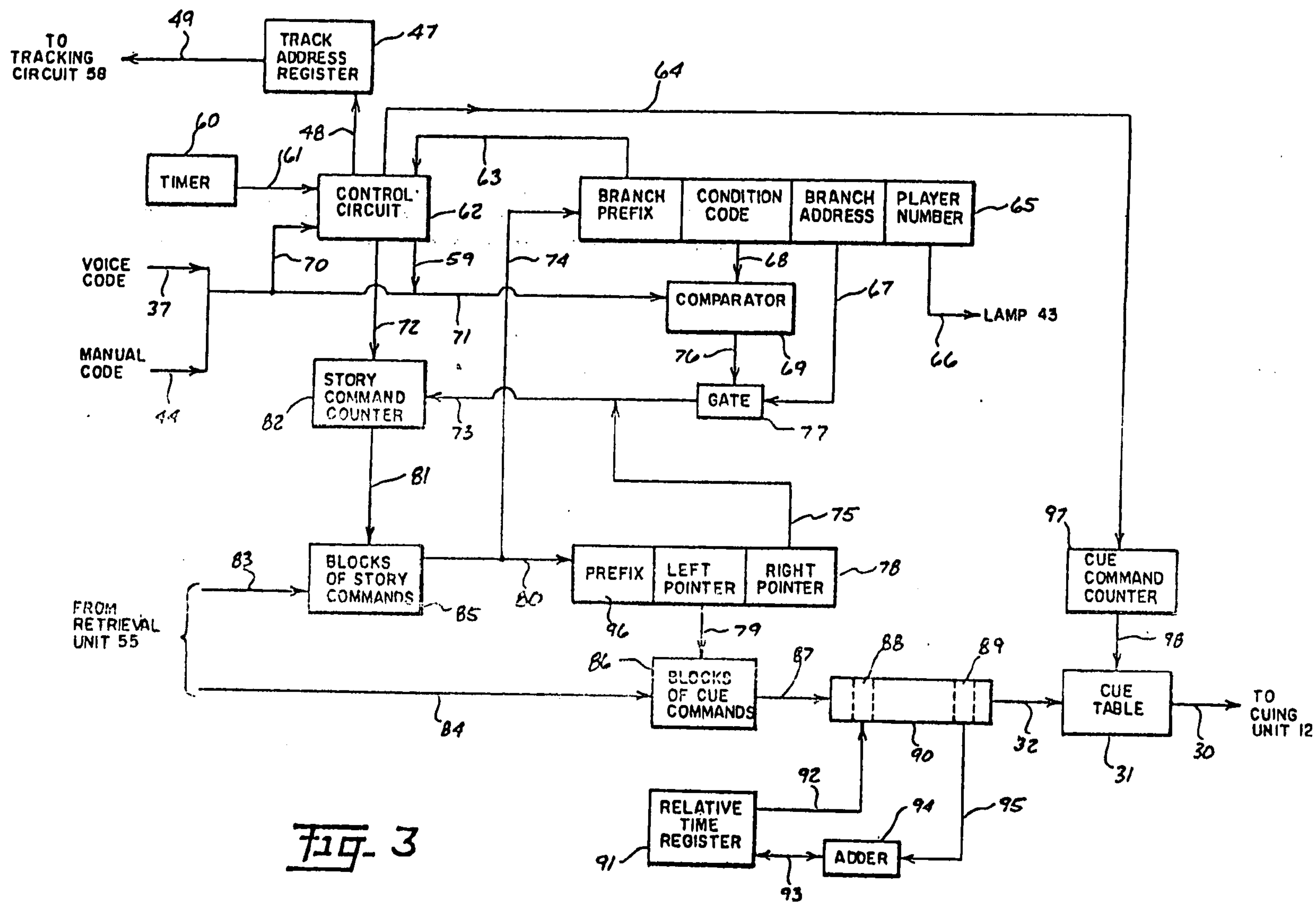

TO TRACKING CIRCUIT 58
49
TRACK ADDRESS REGISTER
47
64
48
60
TIMER
61
CONTROL CIRCUIT
62
63
BRANCH PREFIX
CONDITION CODE
BRANCH ADDRESS
PLAYER NUMBER
65
VOICE CODE
37
70
59
74
68
COMPARATOR
67
LAMP 43
66
MANUAL CODE
44
72
71
76
69
STORY COMMAND COUNTER
82
73
GATE
77
81
75
97
BLOCKS OF STORY COMMANDS
83
85
80
PREFIX
LEFT POINTER
RIGHT POINTER
78
96
CUE COMMAND COUNTER
98
FROM RETRIEVAL UNIT 55
79
86
BLOCKS OF CUE COMMANDS
87
88
89
CUE TABLE
TO CUING UNIT 12
84
90
32
31
30
92
94
95
RELATIVE TIME REGISTER
91
ADDER
93

Fig. 3

0016314

4/15

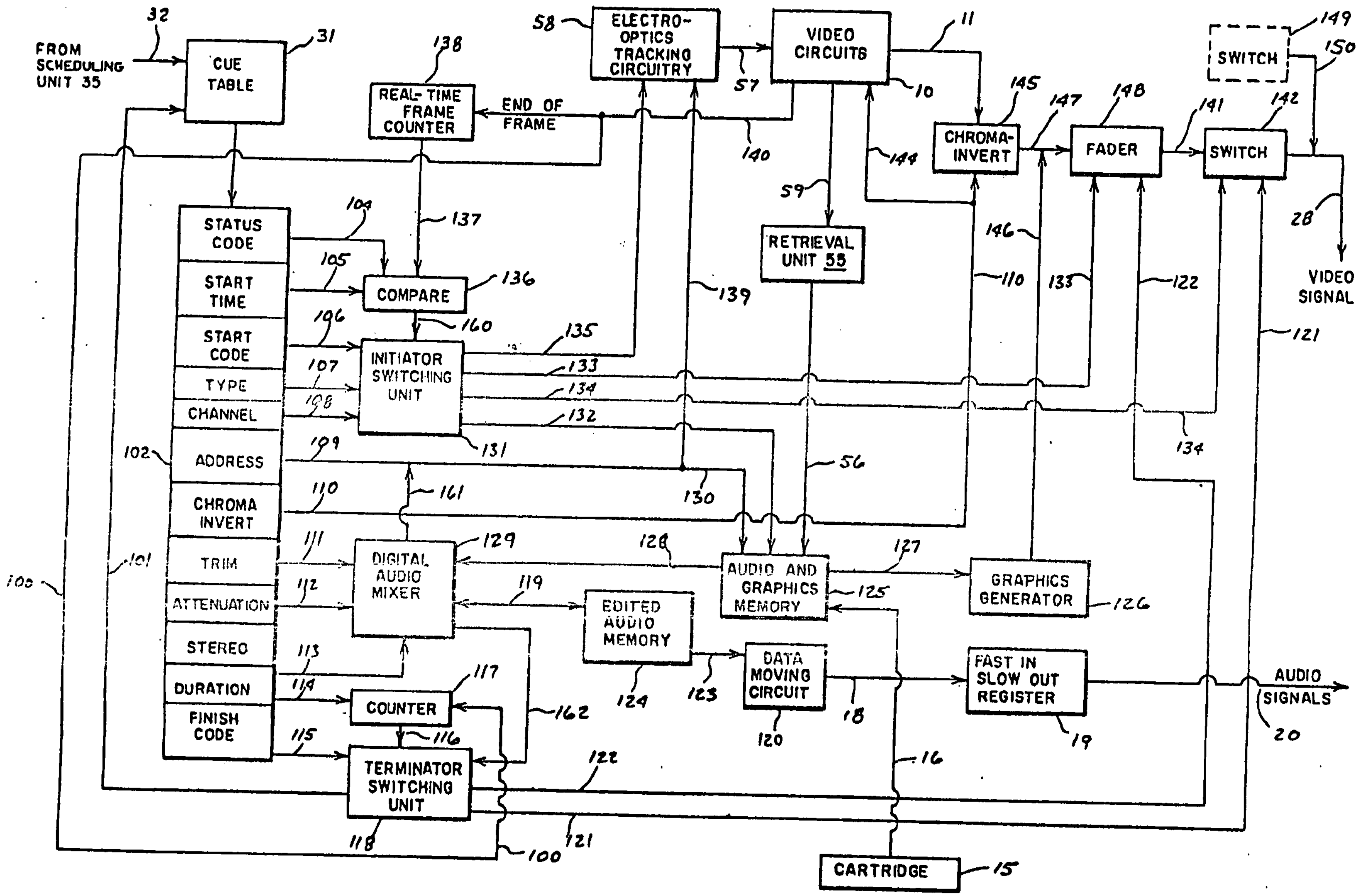


Fig-4

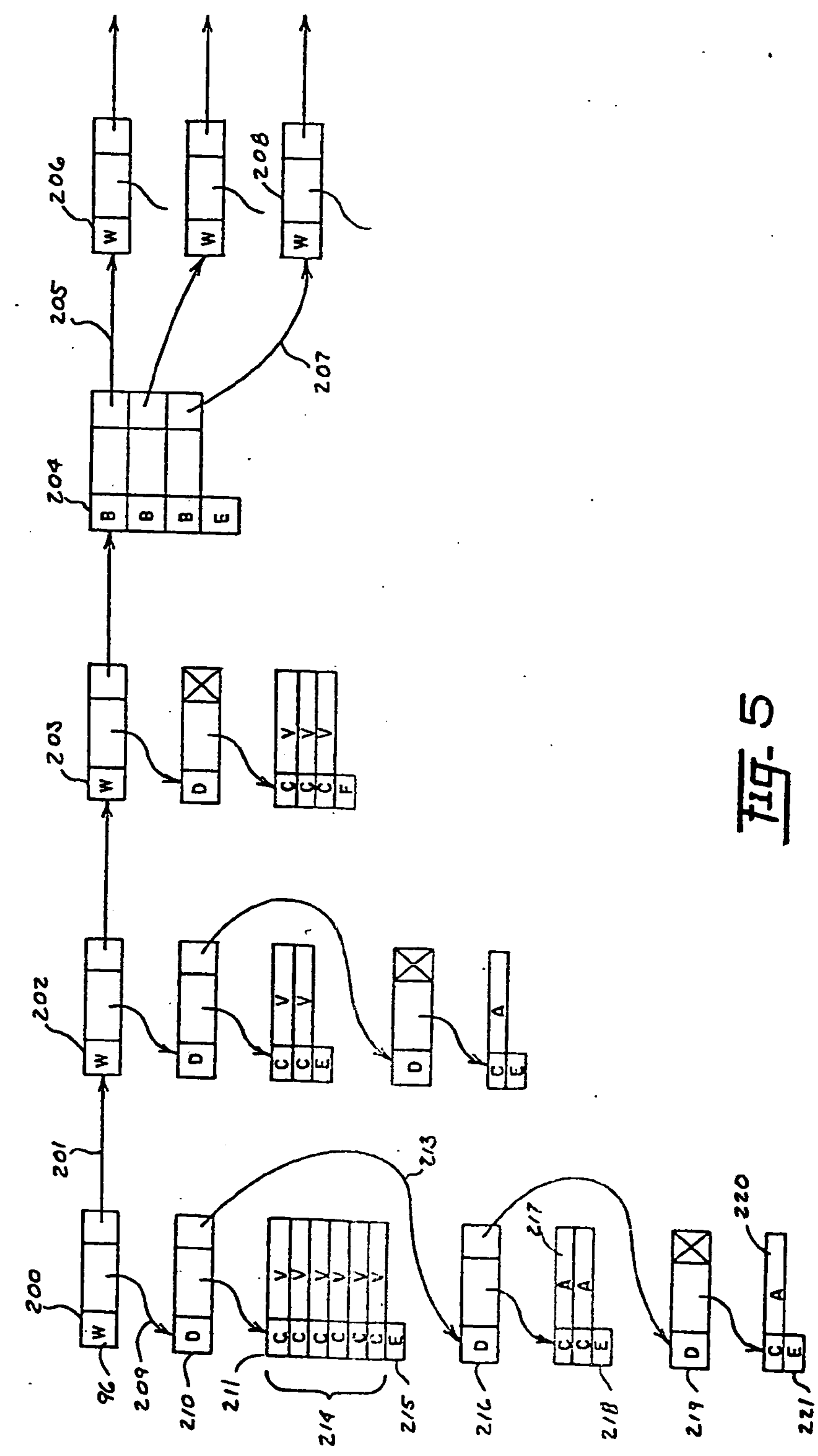

200
201
202
203
204
205
206
207
208
209
210
211
213
214
215
216
217
218
219
220
221
96
W
D
C
V
E
F
B
A

Fig. 5

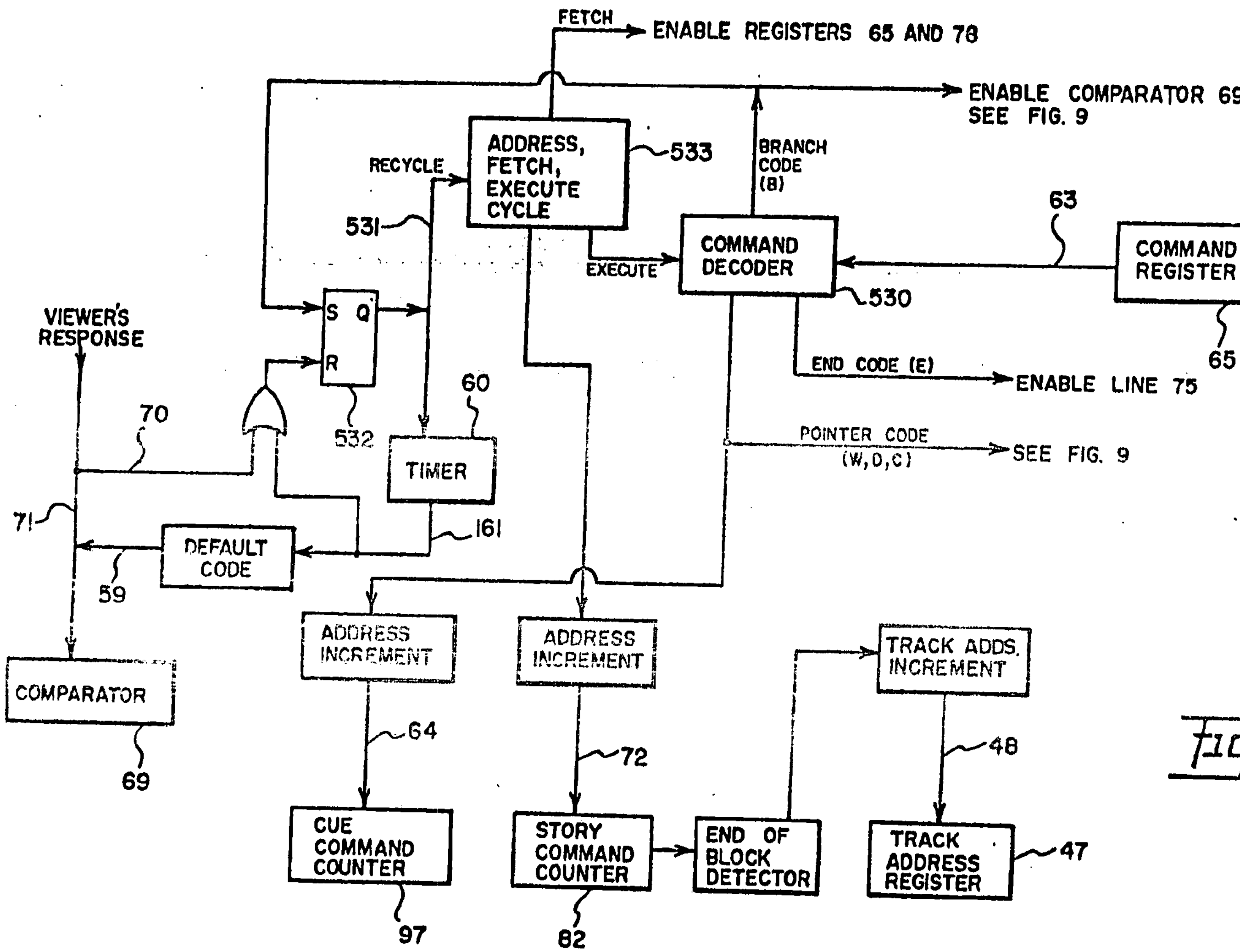

FETCH
ENABLE REGISTERS 65 AND 78
ENABLE COMPARATOR 69
SEE FIG. 9
ADDRESS, FETCH, EXECUTE CYCLE
533
RECYCLE
531
BRANCH CODE (B)
63
COMMAND REGISTER
65
COMMAND DECODER
530
EXECUTE
S Q
R
532
VIEWER'S RESPONSE
END CODE (E)
ENABLE LINE 75
60
TIMER
70
POINTER CODE (W,D,C)
SEE FIG. 9
71
161
DEFAULT CODE
59
ADDRESS INCREMENT
ADDRESS INCREMENT
TRACK ADDS. INCREMENT
COMPARATOR
69
64
72
48
CUE COMMAND COUNTER
97
STORY COMMAND COUNTER
82
END OF BLOCK DETECTOR
TRACK ADDRESS REGISTER
47

Fig. 6

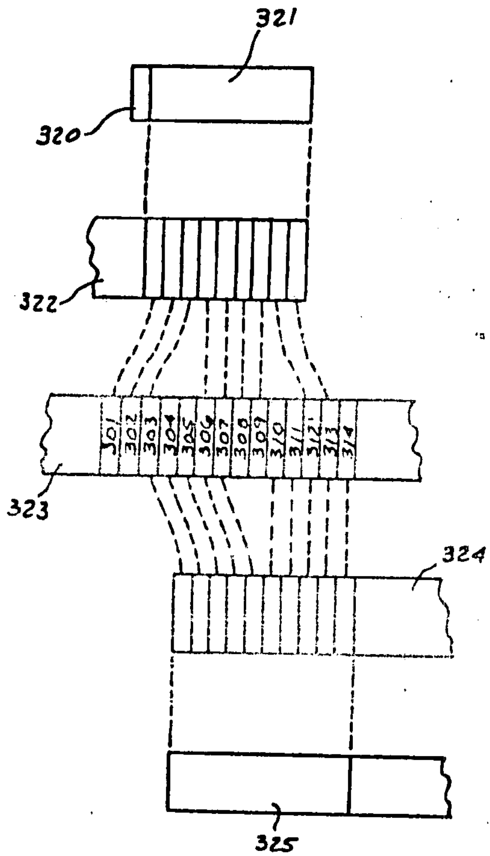

321
320
322
301
302
303
304
305
306
307
308
309
310
311
312
313
314
323
324
325

Fig. 7

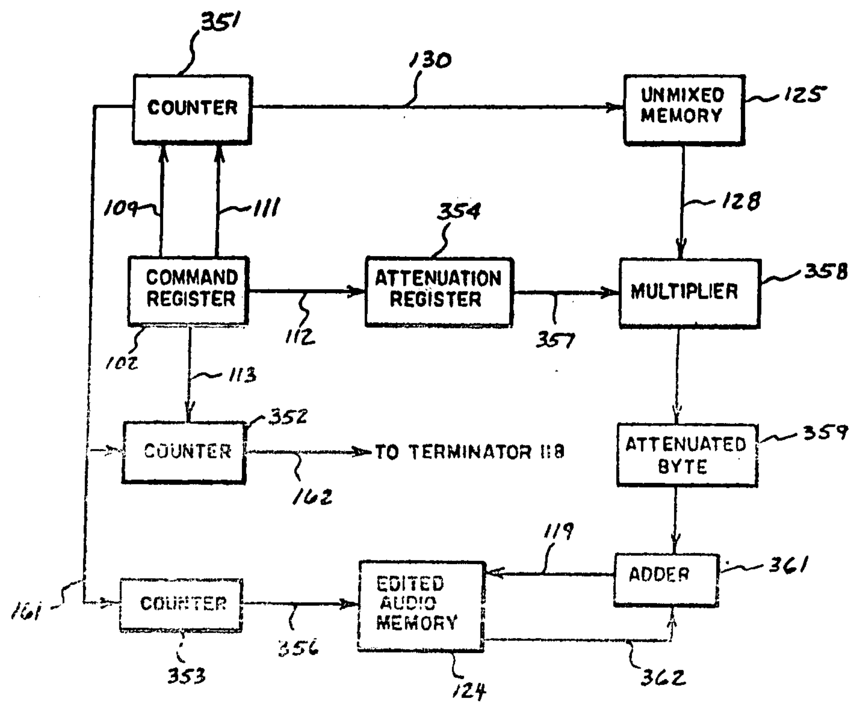

351
COUNTER
130
UNMIXED MEMORY
125
128
109
111
354
COMMAND REGISTER
ATTENUATION REGISTER
MULTIPLIER
358
102
112
357
113
352
COUNTER
TO TERMINATOR 118
ATTENUATED BYTE
359
162
119
ADDER
361
161
COUNTER
EDITED AUDIO MEMORY
356
353
124
362

Fig. 8

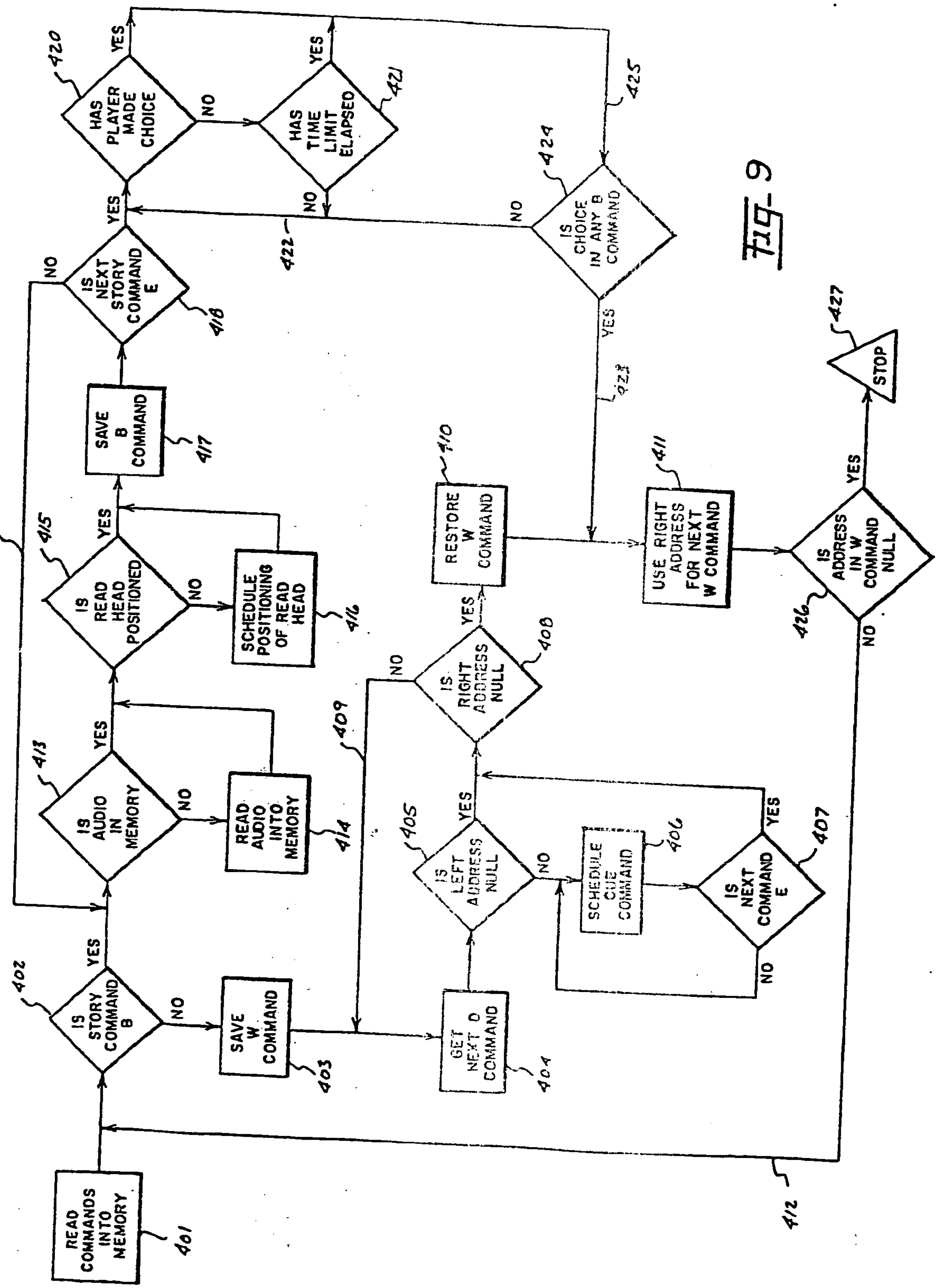
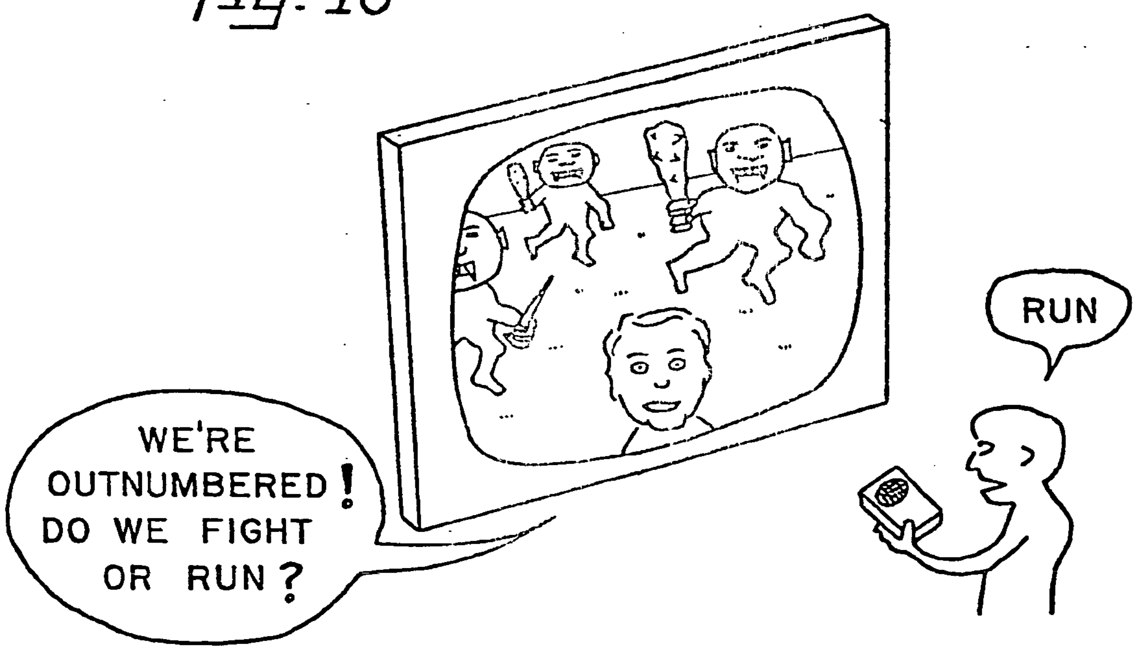

READ COMMANDS INTO MEMORY
401
IS STORY COMMAND B
402
YES
NO
SAVE W COMMAND
403
GET NEXT D COMMAND
404
IS LEFT ADDRESS NULL
405
SCHEDULE CUE COMMAND
406
IS NEXT COMMAND E
407
IS RIGHT ADDRESS NULL
408
409
RESTORE W COMMAND
410
USE RIGHT ADDRESS FOR NEXT W COMMAND
411
412
IS AUDIO IN MEMORY
413
READ AUDIO INTO MEMORY
414
IS READ HEAD POSITIONED
415
SCHEDULE POSITIONING OF READ HEAD
416
SAVE B COMMAND
417
IS NEXT STORY COMMAND E
418
419
HAS PLAYER MADE CHOICE
420
HAS TIME LIMIT ELAPSED
421
422
423
IS CHOICE IN ANY B COMMAND
424
425
IS ADDRESS IN W COMMAND NULL
426
STOP
427

FIG. 9

Fig. 10



10/15

016314

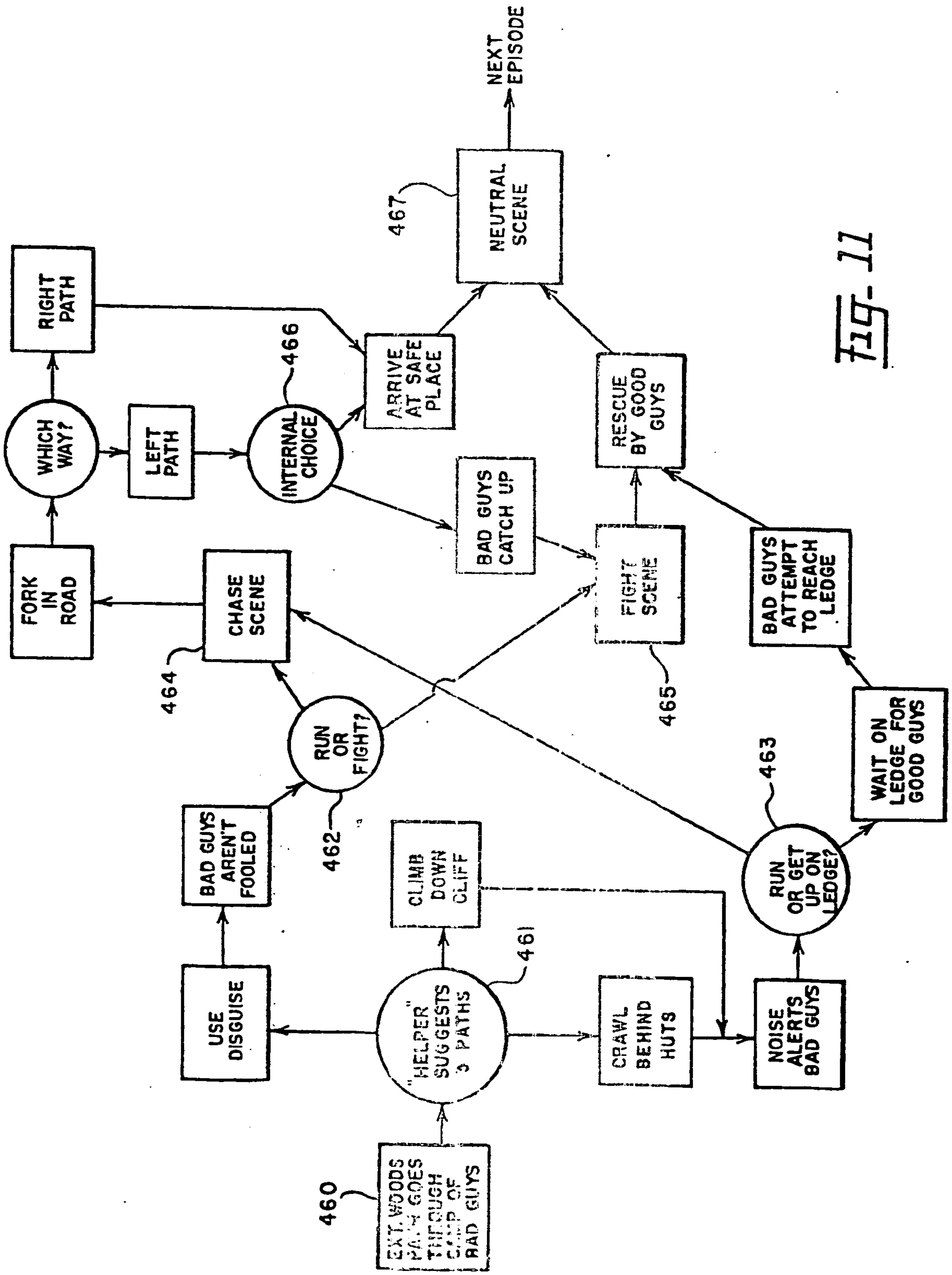


FIG. 11

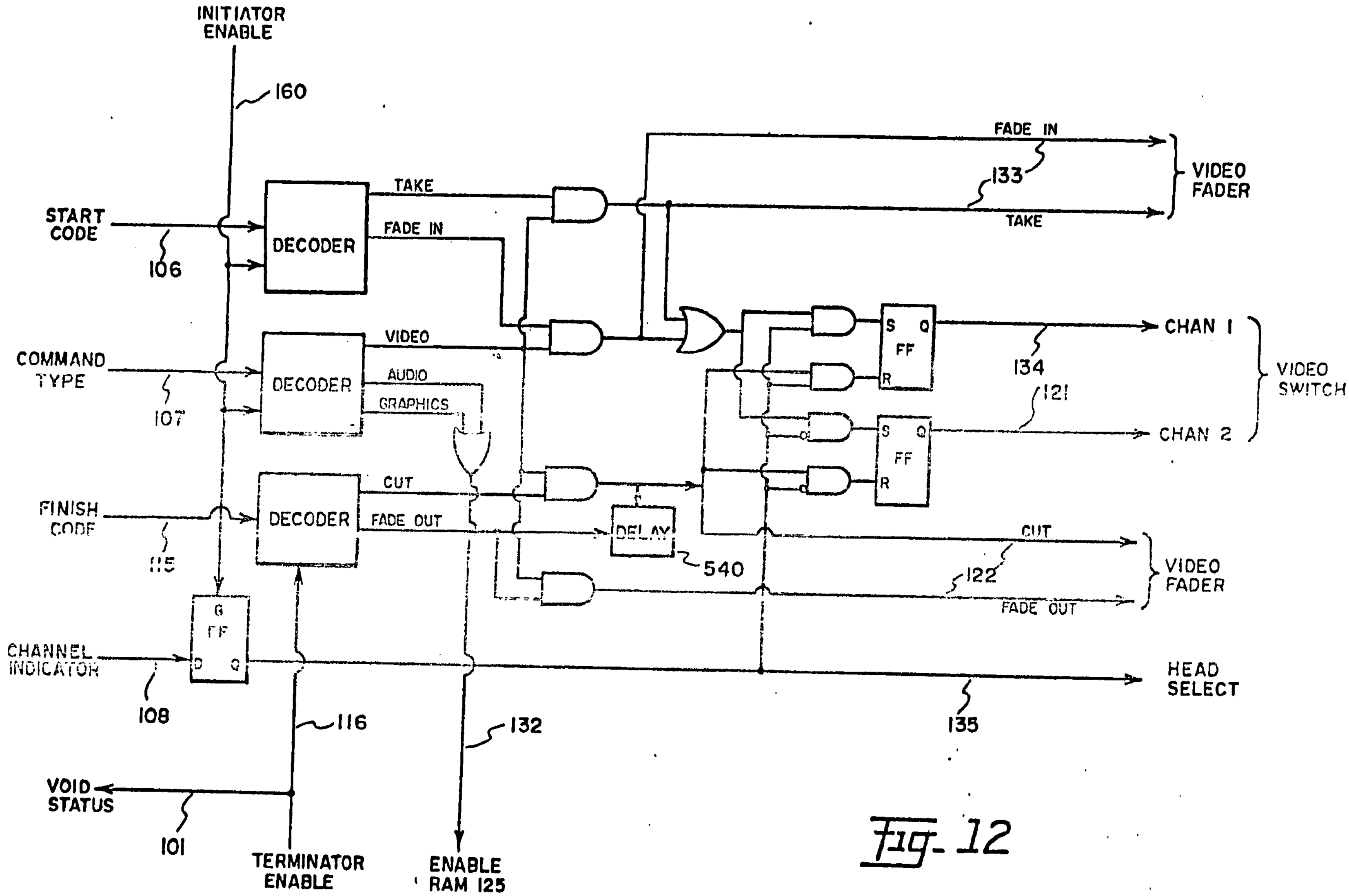

INITIATOR ENABLE
160
START CODE
106
DECODER
TAKE
FADE IN
COMMAND TYPE
107
DECODER
VIDEO
AUDIO
GRAPHICS
FINISH CODE
115
DECODER
CUT
FADE OUT
CHANNEL INDICATOR
108
G
FF
D
Q
VOID STATUS
101
TERMINATOR ENABLE
116
ENABLE RAM 125
132
DELAY
540
S
Q
FF
R
S
Q
FF
R
FADE IN
133
TAKE
VIDEO FADER
CHAN 1
134
121
CHAN 2
VIDEO SWITCH
CUT
122
FADE OUT
VIDEO FADER
HEAD SELECT
135

Fig. 12

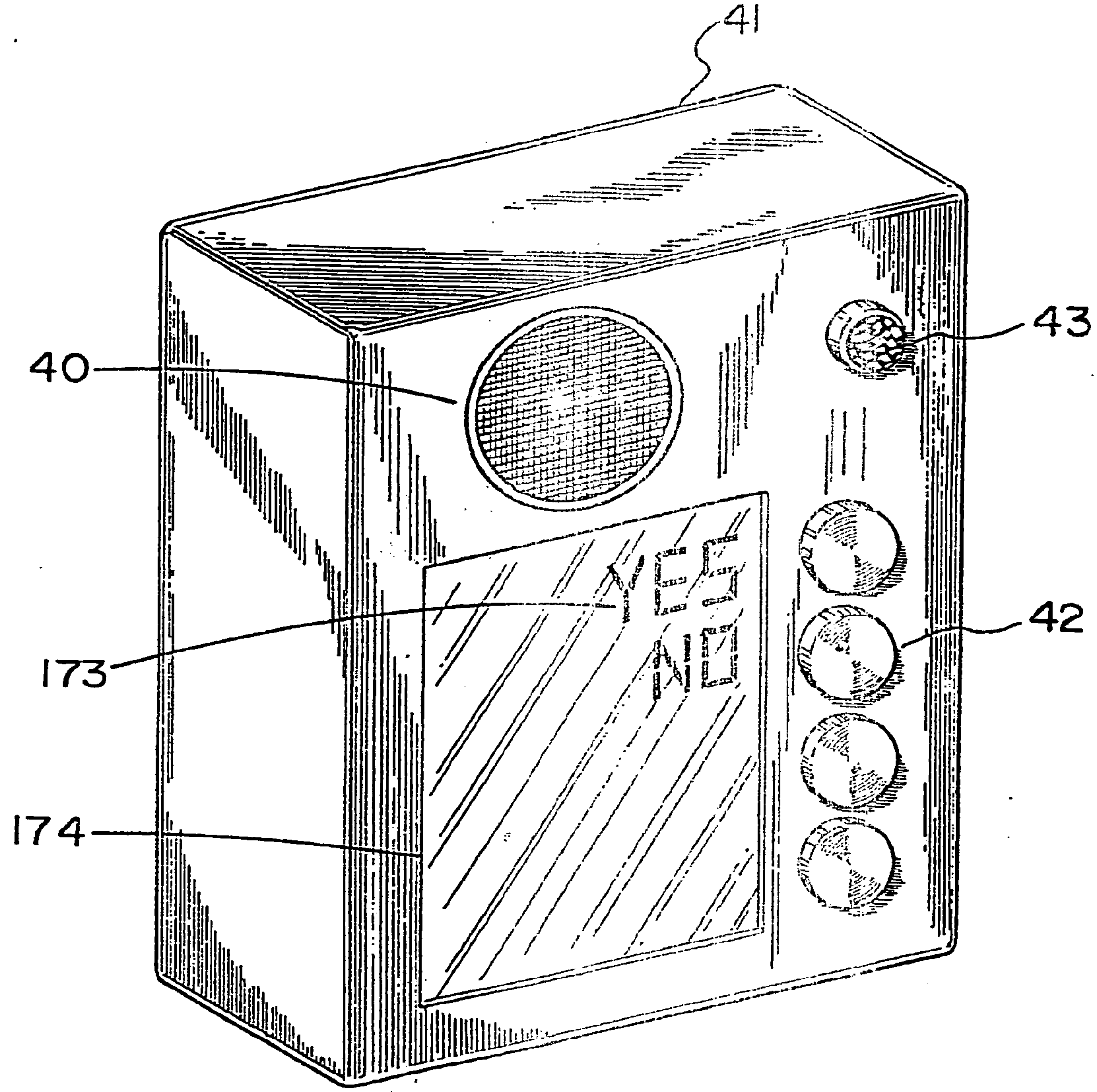


Fig. 13

0016314

13/15

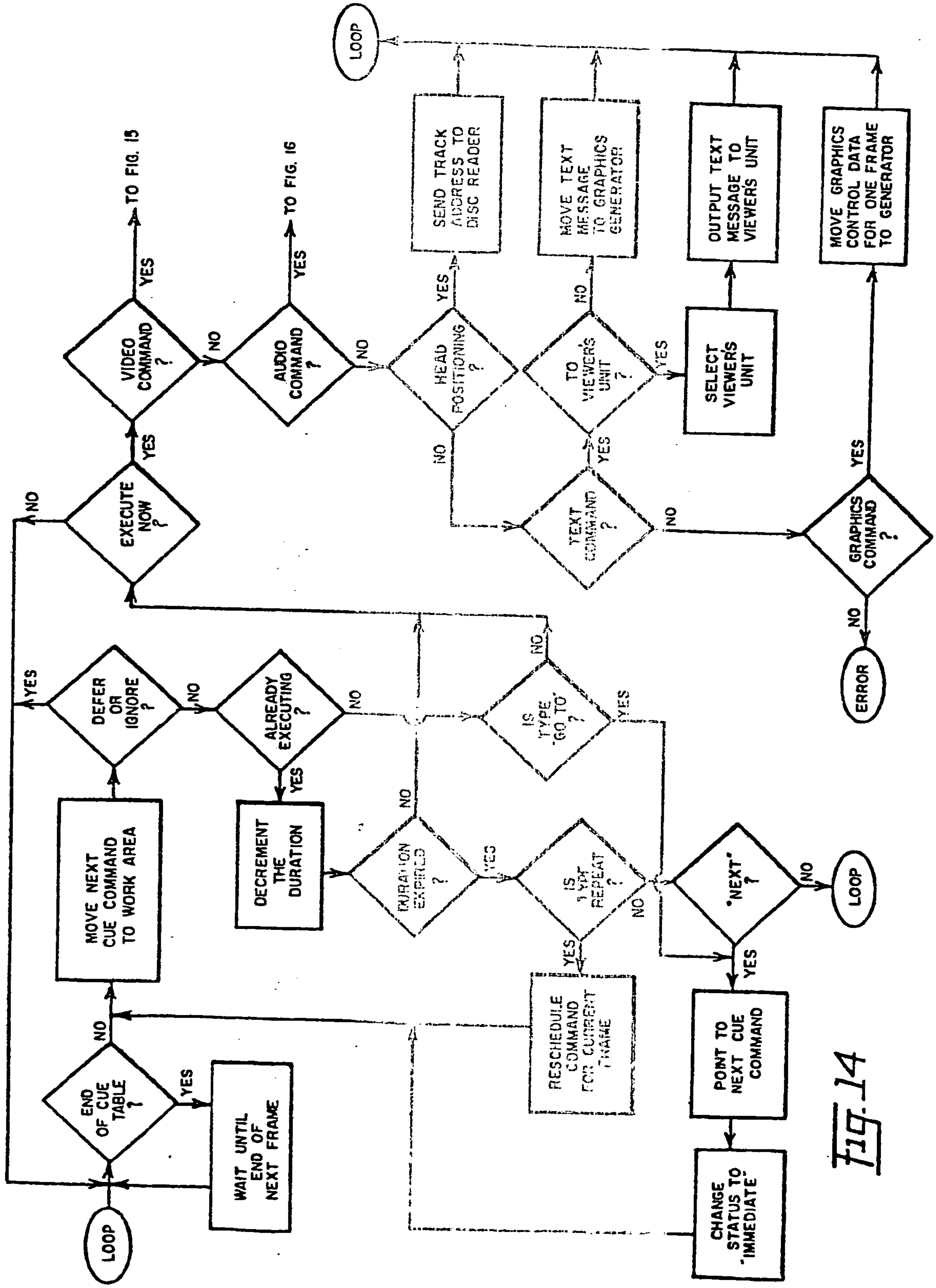


Fig. 14

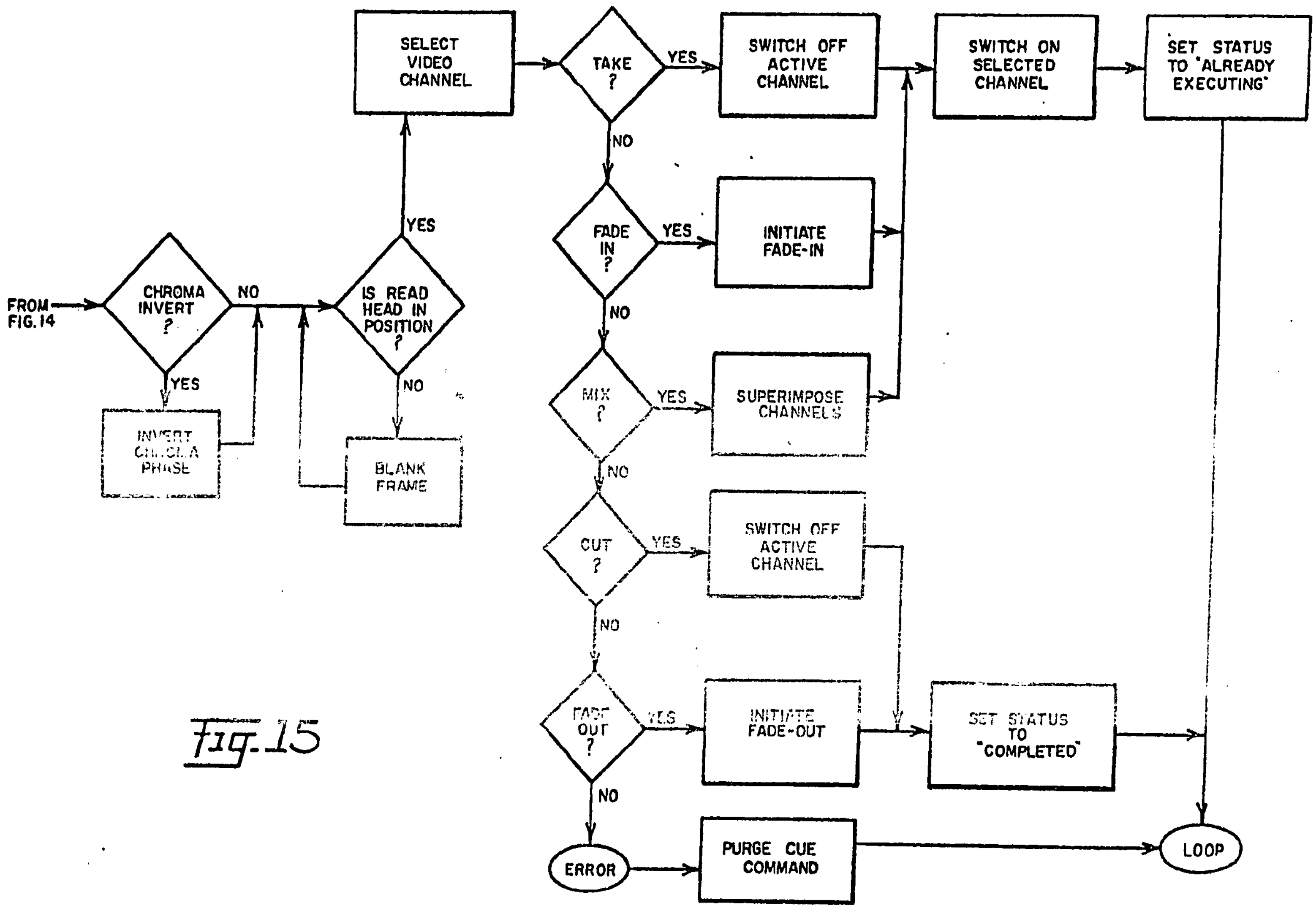

FROM FIG. 14
CHROMA INVERT ?
YES
NO
INVERT CHROMA PHASE
IS READ HEAD IN POSITION ?
YES
NO
BLANK FRAME
SELECT VIDEO CHANNEL
TAKE ?
YES
NO
SWITCH OFF ACTIVE CHANNEL
SWITCH ON SELECTED CHANNEL
SET STATUS TO 'ALREADY EXECUTING'
FADE IN ?
YES
NO
INITIATE FADE-IN
MIX ?
YES
NO
SUPERIMPOSE CHANNELS
CUT ?
YES
NO
SWITCH OFF ACTIVE CHANNEL
FADE OUT ?
YES
NO
INITIATE FADE-OUT
SET STATUS TO "COMPLETED"
ERROR
PURGE CUE COMMAND
LOOP

Fig. 15

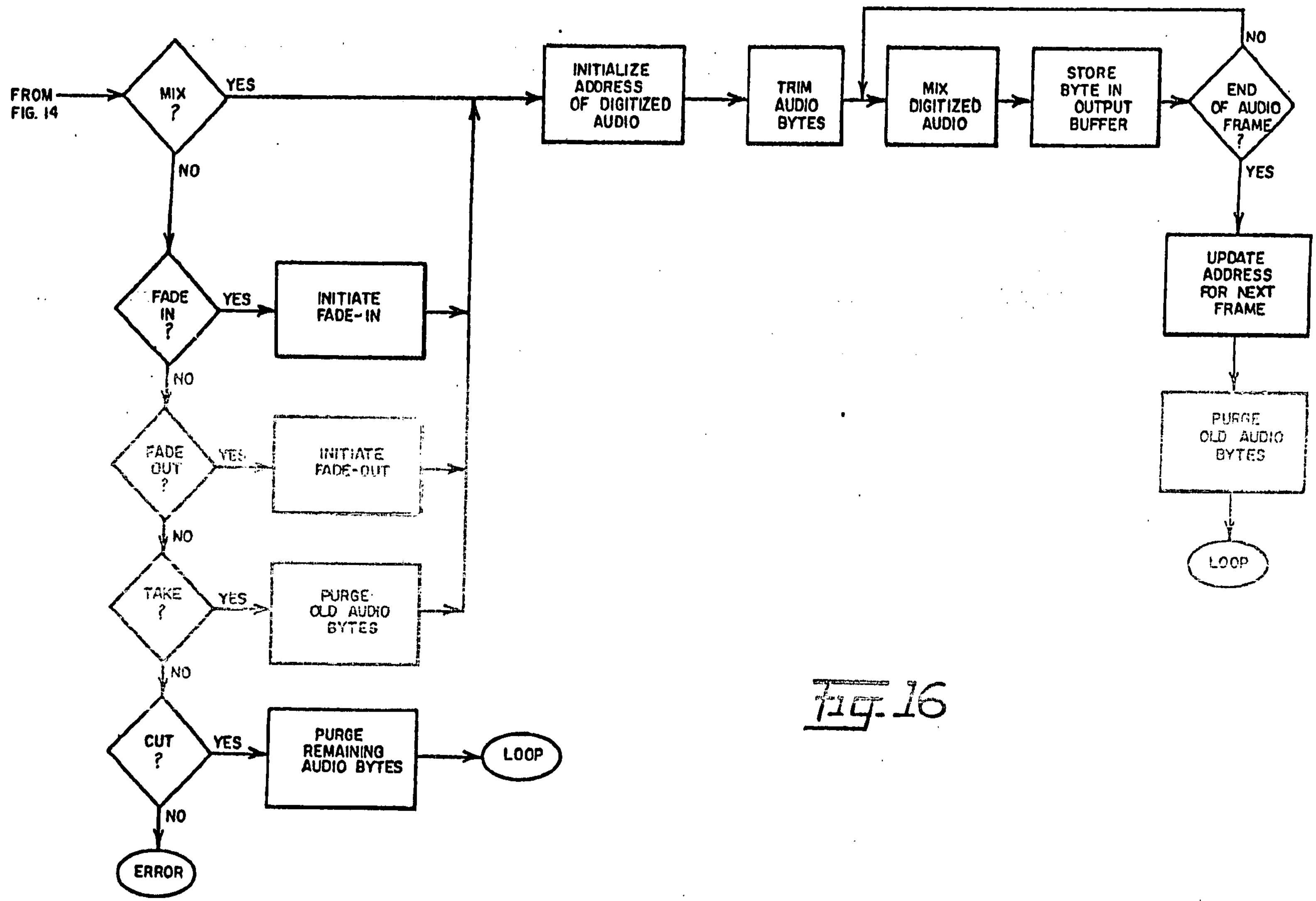

FROM FIG. 14
MIX ?
YES
NO
INITIALIZE ADDRESS OF DIGITIZED AUDIO
TRIM AUDIO BYTES
MIX DIGITIZED AUDIO
STORE BYTE IN OUTPUT BUFFER
END OF AUDIO FRAME ?
NO
YES
UPDATE ADDRESS FOR NEXT FRAME
PURGE OLD AUDIO BYTES
LOOP
FADE IN ?
YES
INITIATE FADE-IN
NO
FADE OUT ?
YES
INITIATE FADE-OUT
NO
TAKE ?
YES
PURGE OLD AUDIO BYTES
NO
CUT ?
YES
PURGE REMAINING AUDIO BYTES
LOOP
NO
ERROR

Fig. 16

0016314

European Patent Office

EUROPEAN SEARCH REPORT

Application number

EP 80 10 0374

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl. 3)
A	JOURNAL OF THE SOCIETY OF MOTION PICTURE AND TELEVISION ENGINEERS, vol. 79, no. 11, November 1970, New York, US, L.W. WELLS: "Random selection and branching in the motion-picture audio-visual field", pages 983-990 * Page 984, right-hand column, lines 35-62 *	1-14	A 63 F 9/22 G 09 B 7/04 H 04 N 5/00
	--		
	US - A - 4 060 915 (CONWAY) * Column 4, line 28 to column 5, line 13 *	1-14	TECHNICAL FIELDS SEARCHED (Int.Cl. 3)
	--		A 63 F 9/22 9/18 H 04 N 5/00 9/00 G 09 B 7/04 7/08 7/12 G 06 F 3/16 G 10 L 1/04
	SCIENTIFIC AMERICAN, vol. 215, no. 2, 1966, New York, US, P. SUPPES: "The uses of computers in education", pages 207-220 * Page 218, caption to figure; page 219, left-hand column, line 28 to center column, line 37 *	1-14	
	--		
	US - A - 3 883 850 (MARTIN) * Column 2, line 45 to column 3, line 17 *	15,16	CATEGORY OF CITED DOCUMENTS X: particularly relevant A: technological background O: non-written disclosure P: intermediate document T: theory or principle underlying the invention E: conflicting application D: document cited in the application L: citation for other reasons
	----		
X	The present search report has been drawn up for all claims		&: member of the same patent family, corresponding document

Place of search	Date of completion of the search	Examiner
The Hague	16-05-...	CRISTOL